AD-A098 230 IOWA UNIV IOWA CITY DIV OF EDUCATIONAL PSYCHOLOGY F/G 12/1
MULTIATTRIBUTE FIXED-STATE UTILITY ASSESSMENT.(U)
MAR 81 D L LIBBY, M R NOVICK N00014-77-C-0428
UNCLASSIFIED TR-81-4 NL

1 OF 3
AD A
098230

A

82

LEVEL II

(12)

Technical Report 81-4

March, 1981

AD A098230

MULTIATTRIBUTE FIXED-STATE UTILITY ASSESSMENT

David L. Libby

The University of Iowa

REPORT PREPARED UNDER OFFICE OF NAVAL RESEARCH
CONTRACT #00014-77-C-0428
AND NATIONAL SCIENCE FOUNDATION
GRANT #SED 80-06357

MELVIN R. NOVICK, PRINCIPAL INVESTIGATOR
UNIVERSITY OF IOWA
IOWA CITY, IOWA

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
Technical Report 81-4

2. GOVT ACCESSION NO.
AD-A098230

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
"Multiattribute fixed-state utility assessment"

5. TYPE OF REPORT & PERIOD COVERED
Technical Report.
7/1/80 - 12/31/80

6. PERFORMING ORG. REPORT NUMBER
Technical Report 81-4

7. AUTHOR(s)
David L. Libby

8. CONTRACT OR GRANT NUMBER(s)
N00014-77-C-0428

9. PERFORMING ORGANIZATION NAME AND ADDRESS
University of Iowa
Division of Educational Psychology
Iowa City, Iowa 52242

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
6115 3N; RR 042-04; RR042-04-01; NR 150-404

11. CONTROLLING OFFICE NAME AND ADDRESS
Personnel and Training Research Programs
Office of Naval Research (Code 458)
Arlington, VA 22217

12. REPORT DATE
March ~~27,~~ 1981

13. NUMBER OF PAGES
189

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
utility, multiple attribute, multivariate probability, conjoint measurement, multidimensional scaling, utility assessment, coherence, expected utility, conditional utility

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
This paper is concerned with the assessment of utility functions for multiple attributes. The literature on utility assessment, multidimensional scaling, conjoint measurement, and conditional expected utility theory is reviewed. A procedure for utility assessment is presented using as models cumulative multivariate distribution functions.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

March 23, 1981

MULTIATTRIBUTE FIXED-STATE UTILITY ASSESSMENT

By

David L. Libby

The University of Iowa

Acknowledgements

This research was supported, in part by the Office of Naval Research under contract N00014-77-C-0428, Melvin R. Novick, Principal Investigator, and in part by the National Science Foundation under Grant #SED 80-06357, Melvin R. Novick, Principal Investigator. Any opinions, findings, conclusions, or recommendations expressed herein are those of the author and do not necessarily reflect the views of the supporting agencies.

Table of Contents

Chapter I. INTRODUCTION

Chapter II. UTILITY ELICITATION

Chapter III. MULTIDIMENSIONAL SCALING

Chapter IV. CONJOINT MEASUREMENT

Chapter V. POLYNOMIAL UTILITY ASSESSMENT

Chapter VI CONDITIONAL EXPECTED UTILITY ASSESSMENT

Chapter VII MULTIATTRIBUTE UTILITY ASSESSMENT

Appendix A The Generalized Beta and F Distributions

Appendix B. Derivatives

Chapter I. INTRODUCTION

Section 1. Overview

In this paper, we develop a method for assessment of utility functions for multiple criteria. Suppose that there is a multidimensional outcome variable, which we shall call $\underline{\theta}$, and a set of acts on which the outcome variable $\underline{\theta}$ depends. In the decision-making procedure, we shall want to choose among the acts when the value of $\underline{\theta}$ is uncertain. In the Bayesian approach to decision making, the problem has two components: a probability function over the variable $\underline{\theta}$ for each possible act, in this case a multivariate distribution; and a utility function over $\underline{\theta}$ for each person. The Bayes rule is to select the act with the highest expected utility with respect to the probability function.

The form of the probability function has been extensively discussed in the literature and does not receive primary attention in this paper. It is, however, relevant to some degree in that, because of considerations of simplicity in the computation of expected utilities, the form of the probability distribution may constrain the form of the derived utility function to the class of conjugate distributions (Novick and Lindley, 1978). It has also been shown (Novick, 1980) that in the multiattribute situation, the relationship between the joint utility and the marginal utilities depends on the (conditional) probability function. Thus the family of conjugate distributions is convenient here as well. With adequate computer-assisted numerical analysis techniques, however, nonconjugate functional forms can easily be used.

The form of the utility function also has received much attention in the literature. The primary references are provided as they are relevant in the presentation of the background material in succeeding sections of this chapter and in succeeding chapters.

The emphasis in this paper is on the form of the utility function. While there has been much research concerning the form of such a function, the results have been insufficiently general for many applications. Much of the discussion has been limited to a unidimensional attribute (e.g., Friedman and Savage, 1948; Mosteller and Nogee, 1951; Pratt, Raiffa, and Schlaiffer, 1965; Keeney and Raiffa, 1976; Novick and Lindley, 1979). Many of the applications of decision theory, particularly in education, however, involve paradigms in which the variable of interest is multidimensional.

The literature that is concerned with multiattributed decision problems is also limited to some degree (Green and Wind, 1973; Keeney and Raiffa, 1976). Most of the research has assumed special forms of the utility function such as an additive or a multiplicative function of the components, or other nearly as restrictive assumptions. While these forms may be appropriate and useful under certain conditions (and these conditions have been well specified), there are many situations in which the necessary conditions fail. Additional research (Fishburn, 1973b) has provided an alternative formulation of the multiplicative utility function, involving sums and products of the marginal functions, yet still imposes conditions that place undesirable constraints on the form of the utility function. For example, in Fishburn's formulation, conditional utility functions (those holding the levels of all but one factor fixed) are strictly linear.

This paper proposes a procedure for the assessment of a multiattribute utility function with a less restrictive functional form. The proposed procedure has three phases: the elicitation phase, in which data are collected; the coherence phase, in which the data and intermediate model estimates of utility are checked for consistency; and the modelling phase, in which the parameters of the proposed model for the utilities are estimated.

The basis of the elicitation phase is the work by Novick and Lindley (1979) on the assessment of utilities for a single attribute. The work of these authors is also used in the coherence phase. The coherence phase incorporates ideas from the field of scaling, particularly multidimensional scaling, as well. For the modelling phase, work from the field of conjoint measurement, as well as from multidimensional scaling, is explored. The theories of conditional expected utility are used to create a marginal and conditional structure for the multiattribute utility function. Finally, the marginal and conditional utility functions are modelled by cumulative distribution functions.

Section 2. The Elicitation Phase

Utilities cannot, in general, be assessed directly; we perceive them only through their impact on decisions. Behavioral data, from which the utilities may be inferred, must be collected. In any procedure to determine utilities, one can distinguish between a data collection phase and an estimation phase. It is this collection of data that we shall call the elicitation phase. Although the choice of estimation procedure is not independent of the choice of the elicitation procedure, in that the method of estimation depends on the type of data that is collected, and although the two phases may be interleaved in an iterative fashion, the two procedures may be construed as distinct phases of the overall procedure.

Much of the research in the determination of utilities does not make the distinction between the elicitation and estimation phases clearly. In fact, some of the advocated procedures are simply called "utility assessment procedures," combining the data collection phase and the parameter estimation phase as one topic (e.g., Novick and Lindley, 1979; Humphreys and Wisuhda, 1979). In a unidimensional utility assessment procedure, this blurring of the component phases is not very serious, because of the simplicity of the procedure. In a multidimensional utility assessment procedure, however, this blurring causes conceptual difficulties because of the greater complexity of the multidimensional procedure. In addition, the modularization adds flexibility in the construction of an appropriate procedure for a given application.

Even in the unidimensional case one may distinquish among elicitation techniques, each of which is appropriate for a given estimation technique, and among estimation techniques, each of which is applicable to the same type of data and elicitation technique. The various utility assessment procedures of Novick, et alia (Novick and Lindley, 1979; Novick, Chuang, and DeKeyrel, 1979; Novick, Hamer, Libby, Chen, and Woodworth, 1980; Novick, Turner, and Novick, 1981), the Standard Least Squares procedure, the Regional Coherence procedure, and the Local Coherence procedure, indicate this flexibility.

This paper does not investigate the various elicitation techniques. The purpose of this paper is to assume the elicitation techniques as developed by Novick, et alia, for the unidimensional case and extend it to a multidimensional setting.

In Chapter II of this paper, the Novick and Lindley procedure is discussed in detail. The chapter begins with a brief background discussion of the von Neumann and Morgenstern axiomatization of utility (von Neumann and Morgenstern, 1947) and the notion of expected utility (Savage, 1954). The assessment procedure of Novick and Lindley is based on these two concepts.

Following this introduction is a general discussion of the procedure, called the fixed-state utility assessment. Briefly, the procedure is based on the concept of greatest expected utility, that individuals choose an act or outcome, among several, having the highest expected utility. Thus, to determine utilities for a variable θ an individual is presented with a choice between one outcome, say θ_j, for sure and a gamble involving two other outcomes, say θ_i and θ_k. The outcomes are selected such that θ_i is not preferred to θ_j which is not preferred to θ_k.

It can be shown that there exists a probability p such that the individual will have no preference for either the for-sure outcome or the gamble over the other. This probability is called the "indifference probability," since the individual is indifferent between the two choices.

Following this is a general discussion contrasting the fixed-state utility assessment procedure to an alternative procedure, called fixed-probability utility assessment. In one paradigm of this procedure, the individual is presented with a choice between a for-sure outcome and a gamble involving two other outcomes. The probability of the gamble is fixed, however, as are the two outcomes involved in the gamble, and the individual is asked to specify the outcome of the for-sure alternative. This is then called the "certainty equivalent," since it is a certain outcome that has the same expected utility as the gamble. A review of the literature that criticizes both these procedures is also presented.

Following these discussions is a detailed presentation of the least-squares estimation procedure advocated by Novick and Lindley and implemented in the Computer-Assisted Data Analysis (CADA) Monitor (Isaacs and Novick, 1978; Novick, Hamer, Libby, Chen, and Woodworth, 1980). Criticisms of the 1978 implementation by this author are presented, which led to the enhanced, 1980 implementation. Additional enhancements are suggested in this chapter as well.

Section 3. The Coherence Phase

Just as the procedure to assess utilities may be decomposed into an elicitation phase and an estimation phase, so may the estimation phase be decomposed into a coherence phase and a modelling phase. In the coherence phase, the data and the intermediate estimates of the utilities from the modelling phase are checked for consistency. The responses from the user may appear to be reasonable at face value, but they may imply relationships among the attributes that are unacceptable. There are two concepts of coherence that are explored in this paper. One is from the work of Novick and Lindley and is discussed in Chapter II. The other is related to the field of scaling. Both concepts of coherence depend upon the assumptions that the decision maker is willing to accept about the data.

The Novick-Lindley concept of coherence (Novick and Lindley, 1979; Novick, Chuang, and DeKeyrel, 1979; Novick, Turner, and Novick, 1981) involves judgments by the decision maker. Implications of the data and the selected models are presented and the decision maker must judge whether the implications are acceptable. If they are not, additional information must be elicited to resolve the incoherence. In the paradigm of a single decision maker used by Novick and Lindley, the decicion maker is asked to reconsider some of the original judgments (i.e., the data). Alternatively, gathering more data may be more practical, particularly in a problem involving many sources of the data. This latter approach is often taken in classical statistics, for example, in discriminant function analysis with cross-validation studies.

In scaling, one is concerned about coherent judgment as well. It is assumed that the scale of measurement as observed may not be the scale of measurement with which decisions should be made. For example, in our decision, we might like to assume that equal intervals have equal meaning everywhere along the scale. We might be unwilling to make this assumption about our elicited data. Therefore, the elicited data are transformed so that they may be manipulated by the operations of arithmetic in the modelling phase. The level of measurement of the elicited data may not be strong enough to allow the usual operations of arithmetic to be meaningful. The theory of measurement has defined four categories: nominal, ordinal, interval, and ratio (Stevens, 1946; Coombs, 1950; Coombs, 1951; Siegel, 1956a; Siegel, 1956b).

If the data collected form a nominal scale of measurement, i.e., form a classification only, the average of two data values has no sensible interpretation. There is no sense of order between data values, much less the sense of distance that is necessary to interprete the idea of an average. The only relation that is defined for a nominal scale of measurement is equivalence. Either an object is in the same equivalence class as another or it is not.

In an ordinal level of measurement, not only is there a sense of equivalence, but there is an ordering relation as well. The ordering relation may be conceived as "greater than", "preferred to", "lower than", "lighter than", et cetera. There is no sense of distance. One can say that an object is "preferred to" another, but one cannot say by how much. This level of measurement merely ranks the objects by some property. Any monotone transformation of the scale provides an equivalent scale.

With an interval level of measurement, the zero point and the unit of measurement are arbitrary, but constant. Intervals between objects are known and the ratio of any two intervals does not depend on the unit of measurement or the zero point. Any linear transformation of the data results in an equivalent scale.

In a ratio level of measurement, the scale has all the properties of an interval scale plus a defined zero point. This implies that the ratio of any two values has a meaningful interpretation; one can judge that one object is twice another. Any transformation that leaves the ratio of two values constant defines an equivalent scale. Thus, the unit of measurement is arbitrary.

Many of the advocated utility assessment procedures ignore this distinction or omit the scaling entirely. For example, in the Novick and Lindley fixed-state utility assessment procedure, it can be shown that the indifference probabilities are assumed to be from a ratio level, or scale, of measurement (Edwards, 1965). This is due to the probabilistic interpretation and the estimation procedure. This is a strong assumption to make, particularly if the procedure is used by those who are unfamiliar with the gambling context of the Novick and Lindley procedure or the notion of probability.

This assumption may be relaxed to some extent. To this end, two alternatives are available. One is to use a different data collection procedure in which the natural level of measurement is less restrictive (e.g., interval or ordinal, instead of ratio). This alternative has been used with success (Kruskal, 1965; Green and Wind, 1973). The other is to assume a less restrictive level of measurement for the indifference probabilities and to apply a scaling transformation to obtain the ratio level of measurement required by the estimation procedure.

To provide the necessary background material for the scaling, the literature of multidimensional scaling is reviewed in Chapter III (Eckart and G. Young, 1936; G. Young and Householder, 1938; Torgerson, 1952; Kruskal, 1964; Shepard, 1966; Guttman, 1968; Carroll and Chang, 1970; F. W. Young, de Leeuw, and Takane, 1976a). The history of multidimensional scaling is traced from Eckart and G. Young through the theories of data and modelling of F. W. Young, et alia.

Following this, the theory of data of F. W. Young is presented in detail. The theory is based on two concepts: level of measurement and process of measurement. The level of measurement is concerned with comparisons between categories of objects. The different levels are nominal, ordinal, interval, and ratio. These levels have been discussed above. The process of measurement is concerned with comparisons within categories. The two processes of measurement are discrete and continuous. In a discrete process of measurement, all objects in the same category are assumed to have the same value on the scale. If two objects have different values, they are in different categories. In a continuous process of measurement, the categories are defined by intervals of values, and all objects having values in a given interval belong to the same category.

The utility assessment procedure of Novick and Lindley includes some scaling of the elicited data. These authors argue that equal intervals, in the elicited data, near the middle of the scale and near the tails of the scale are not equally meaningful (Novick and Lindley, 1979). Therefore, they suggest using the logodds transformation to balance the differences. This may be viewed as a fixed, non-parametric scaling transformation.

The main purpose of reviewing this literature is to show the similiarities between the research in scaling and in utility assessment.

Section 4. The Modelling Phase

Once we have scaled the data that we have collected, we must proceed to the stage of estimating the utilities. We do this by proposing a model relating the utilities to the data and estimating the parameters of the model by some means such as the method of least squares. Much of the multidimensional scaling literature is useful in investigating the modelling phase; thus, this part of the utility assessment procedure is discussed in Chapter III as well.

In the literature on utility estimation, however, most of the modelling is discussed in terms of conjoint measurement. Background material for conjoint measurement is presented in Chapter IV. Two axiomatizations are discussed (Luce and Tukey, 1964; Roskies, 1965). Several applications conforming to these axiomatizations are presented (Fishburn undated; Keeney and Raiffa, 1976; Kruskal, 1965; Green and Wind, 1973; Humphreys and Wisuhda, 1979). These applications depend upon certain independence conditions that simplify the form of the utility function.

One of these conditions is called "additive independence" (Keeney and Raiffa, 1976). This condition holds if and only if preferences among gambles depend only on the marginal probability distributions of the gambles. When the condition holds, the utility function is additive in the components.

Another of these conditions is call "mutual utility independence" (Keeney and Raiffa, 1976). Under this condition, preferences among gambles involving each attribute do not depend on the levels of the other attributes. If this condition is satisfied, then the utility function is multilinear; i.e., it is the sum of uniattribute utility functions and cross-products of uniattribute utility functions.

The points to be made here are that the conditions may not be satisfied and that the resulting functional limitations may not be appropriate. We must be careful to investigate the validity of the conditions and the meaningfulness of the functional forms. It is likely that an additive utility function is not reasonable in most situations. A multilinear utility function is also unreasonable in many situations. In both cases, for example, the conditional utility functions, those of one attribute for fixed values of the others, are linear. While these simplifications are useful under the correct conditions, they are not generally applicable. Methods for estimating more general multiattribute utility functions must be developed.

Two general theories of conjoint measurement are presented (Tversky, 1967a; F. W. Young, 1972). These two theories of polynomial conjoint measurement provide the theoretical foundation for a more general approach to the estimation of multiattribute utility functions. The approach taken in Chapter V of this paper is based on the operational approach of F. W. Young, et alia (de Leeuw, F. W. Young, and Takane, 1976; F. W. Young, de Leeuw, and Takane, 1976a; Takane, F. W. Young, and de Leeuw, 1977; F. W. Young, Takane, and de Leeuw, 1978).

Section 5. *An Intermediate Proposal*

In Chapter V, the various pieces are put together to form one possible procedure for multiattribute utility assessment using theories from conjoint measurement and multidimensional scaling. The indifference probability assessment procedure is expanded to involve gambles among several attributes. The scaling transformations are developed for several assumed levels of measurement. A polynomial model is proposed that should be more flexible than the additive or the multiplicative model.

Briefly, the multiattribute utility assessment procedure consists of the three phases outlined in the previous three sections: an elicitation phase; a coherence phase; and a modelling phase.

In the elicitation phase, choices between one multidimensional outcome for sure and a gamble involving two other multidimensional outcomes are presented. These choices are used to order the outcomes as well as to elicit the indifference probabilities for the choices. Note that if one choice is selected, either the for-sure outcome or the gamble, regardless of the probability of the gamble (i.e., even when the indifference probability is 0 or 1), then we know that the utility of the for-sure outcome is not between the utilities of the outcomes in the gamble. We can then reorder the outcomes and present a new choice.

The number of possible choices is quite large. If there are d dimensions and N_i is the number of points selected along the ith dimension, then there are (N_1*...*N_d choose 3) possible choices. For example, if there are three dimensions and seven points along each, then there are 6,666,891 possible choices. By comparison, if there is only one dimension and seven points along it, there are only 35 possible choices.

It is not practical to obtain indifference probabilities for every possible choice, nor is it necessary. If there are three dimensions and seven points along each, then there are twenty-one marginal utility parameters, ignoring for the moment the scaling parameters. Recalling that two parameters are fixed, at 0 and 1, for each utility marginal function along its dimension, we have only fifteen parameters to estimate.

We could accomplish the same structuring using the usual unidimensional utility assessment procedures. If we again assume that we are interested in a situation with three dimensions and seven outcomes along each dimension, we have 341 utility parameters to estimate (7*7*7-2=341). This approach makes no assumptions about the structure of the model. The savings in parameter estimation using the polynomial model advocated in this chapter are considerable.

The selection of the choices must be made carefully. One could randomly select multidimensional outcomes for the choices, but this method would likely be confusing for the user. Perhaps, the choices could be selected in a pattern, conditional on certain "lines" through the space of the outcomes. For example, one set of choices could involve points along a single dimension, conditional on fixed values for the other dimensions. Another set of choices could involve points along a "diagonal" through two or more dimensions of outcomes, holding the values of other dimensions fixed.

The second phase, coherence, and the third phase, modelling, would be iterative. An initial scaling transformation, an identity transformation, would be applied to the elicited data. The parameters of the model would then be estimated from this transformed data. Using the estimated parameters of the model, predicted data would be calculated. A new scaling transformation would then be estimated using the predicted data and the original data. This iteration would continue until convergence criteria are satisfied.

Within the third phase, there might be some subiteration. The proposed model might be such that the parameters cannot be estimated simultaneously with ease. F. W. Young, et alia, have shown that the method of alternating least squares works well. With this method, blocks of parameters are estimated conditionally on fixed values for the other parameters. These estimates are then taken to be fixed and another block of parameters is estimated.

The model proposed in this chapter is a fourth-order polynomial. The terms could be determined dynamically: if a term accounts for at least a certain percentage of the total variation in the data, then it is included in the model. Note that the parameters of the model are not only the coefficients of the terms in the model. Since this is a fixed-point procedure, the utilities along each of the dimensions are parameters as well. Any interpolation of utilities between specified outcome points is risky, but useful. To facilitate this interpolation, transformations from the attribute space to the marginal utility spaces are also estimated by fourth-order polynomials.

It should also be noted that the above presentation has tacitly implied that the number of dimensions for utilities is the same as the number of dimensions for attributes. This is not a necessary restriction. Multidimensional scaling is commonly used as a dimension-reducing technique. It also commonly assumes that the dimension of the object space is the number of outcomes. We could apply the techniques presented in this paper to a set of outcomes that appear to be unidimensional on the surface, but are composed of several underlying dimensions. The procedure could then be used to estimate utilities on these underlying dimensions.

The procedure could also be used in reverse. Suppose that we measure outcomes along several dimensions, but that some of the dimensions are redundant. We could use the techniques presented here to identify the utility space of lower dimension.

Section 6. Conditional Expected Utility Assessment

The procedure developed in Chapter V has some deficiencies in conceptualization and in parsimony. Thus the theories of conditional expected utility (Luce and Krantz, 1971; Fishburn, 1973c) are explored as an alternative structure for the multiattribute utility function in Chapter VI. The axiom systems of Luce and Krantz (1971) and of Fishburn (1973c) are briefly presented to provide the theoretical background for the proposed assessment procedure.

Drawing on the two axiom systems, the concepts of a marginal utility function and a conditional utility function are developed. The marginal utility function is defined as the conditional expectation of the multiattribute utility function. The conditional utility function is defined as a rescaled utility function over a subset of the space of interest. It is shown that the conditional utility function plays an integral part in the assessment procedure of Novick and Lindley (1979).

Section 7. Multiattribute Utility Assessment

Based on the theoretical background of Chapter VI, an assessment procedure for the marginal and conditional utility functions is adopted from that of Novick and Lindley. Models for the utility functions are chosen to be cumulative distribution functions, as suggested by Novick and Lindley (1978). By selecting convenient cumulative distribution functions, such that the marginal and conditional distributions are nicely related to the parent distribution, the parameters of the overall multiattribute (parent) utility function may be estimated. The normal distribution and the Dirichlet distribution are convenient candidates. A new distribution, developed for this application and called the multivariate generalized beta distribution, is also used.

The procedure will be implemented into the Computer-Assisted Data Analysis (CADA) Monitor (Novick, Hamer, Libby, Chen, and Woodworth, 1980). It will be limited to two-dimensional problems because of computer memory restrictions. For the same reason, the scaling transformations of the coherence phase will be limited to fixed, non-parametric forms. Three forms will be available: an identity transformation, so that the parameters of the utility model are estimated in the metric of the indifference probabilities;

the logodds transformation, as used by Novick and Lindley (1979) and currently implemented in the CADA Monitor; and the arcsine-square-root transformation, because of its accepted use in educational and psychological research (Novick and Jackson, 1974).

Chapter II. UTILITY ELICITATION

Section 1. Introduction

In this paper, we are primarily concerned with the assessment of utilities for an individual. The data consist of repeated measures elicited from that individual. By contrast, if we are estimating utilities for a group, then the source of error of primary concern is across individuals. In the former case, we want to elicit a large enough sample from the individual so that the errors of measurement are reduced. In the latter case, we generally want to sample sufficiently many individuals to accomplish the same reduction in the error of measurement. In both cases, we want a large enough sample of measurements to achieve stability of the estimates.

When we are assessing utilities for a group, we use the frequency of preference of one object over another to measure the relationship of the two objects (Shepard, Romney, and Nerlove, 1972; Green and Wind, 1973) We use the relative frequency of preferences both to determine the ordinal characteristics (i.e., which has greater utility) and to scale the objects in terms of distances (i.e., how much greater utility the one has compared to the other). If one object is preferred over another by a large majority of the sample and a third over a fourth by a small majority of the sample, one would say that the two objects of the first pair were more "distant" from each other than the two objects of the second pair. The strength of the interval relationship is determined by the proportion of the sample that prefers the one object over the other.

When we are assessing utilities for individuals, we do not have the same type of information from which to draw inferences. We have a sample of one, using the same context. We must use the sample of responses from that one individual to estimate the utilities, but the responses must be of a different nature than a simple statement of preference between two objects. Such an elicitation would be rather transparent, since the individual would likely remember the preference stated for previous presentations of the pair of objects, and thus would not provide as much information to us for our efforts as it might appear. The strongest statement we could make from such evidence would concern only the ordinal characteristics of the utilities, and that would not be very strongly supported.

In order to obtain more information, there are several feasible alternatives. One alternative is to present two pairs of objects and to elicit which pair is thought to be further apart (Suppes and Winet, 1955; Mayekawa, 1980). This comparison on an interval scale is appealing as we are are attempting to assess interval-scaled utilities.

Another alternative is to elicit a statement about the magnitude of the interval between two objects. This procedure might be difficult, and thus unreliable, because of the lack of a standard for comparison. To structure this procedure a bit, we might consider three objects, instead of two, and ask how much further the third is from the first than the second is from the first.

This latter procedure captures in essence one interpretation of an elicitation procedure based on the theory of expected utilities (von Neumann and Morgenstern, 1947; Savage, 1954). The next section presents the axiomatizations of utility by von Neumann and Morgenstern and by Savage. In the third section, two classes of procedures based on the theory of expected utility are investigated: fixed-probability utility assessment and fixed-state utility assessment. In the fourth section of the chapter, the fixed-state procedure advocated by Novick and Lindley (1979) is discussed in detail. In the final section, some enhancements for the Novick-Lindley procedure are presented.

Section 2. Expected Utility Axiomatizations

In this section, two axiomatizaticns of expected utility are presented. The axioms of von Neumann and Morgenstern (1947) are listed in Table II.2.1. Those of Savage (1954) are listed in Table II.2.2. These axiomatizations are presented here for completeness and are not discussed in detail.

The axiomatization of von Neumann and Morgenstern was the first concerted axiom system for the concept of expected utility. The concepts may be traced back to the works of Ramsey (1960) and de Finetti (1974). The axiom system of Savage generalizes the system of von Neumann and Morgenstern somewhat and presents it in a more rigorcus fashion.

Section 3. Expected-Utility Assessment Procedures

In this section we present two classes of procedures based on the theory of expected utility. The procedures in one class are called fixed-state procedures, and those in the other class are called fixed-probability procedures.

Table II.2.1

Axiomatization of Utility

(von Neumann and Morgenstern 1947)

We consider a system U of utilities u, v, w, ... In U a relation is given, $u > v$, and for any number a, $(0<a<1)$, an operation

$$au + (1-a)v = w.$$

These concepts satisfy the following axioms:

A $\quad u > v$ is a complete ordering of U.

A:a $\quad$ For any two u, v, one and only one of the three following relations holds:

$$u = v,\ u > v,\ u < v.$$

A:b $\quad u > v$, $v > w$ imply $u > w$.

B $\quad$ Ordering and combining

B:a $\quad u < v$ implies that $u < au + (1-a)v$

B:b $\quad u > v$ implies that $u > au + (1-a)v$

B:c $\quad u < w < v$ implies the existence of an a with

$$au + (1-a)v < w.$$

Table II.2.1

(continued)

B:d $u > w > v$ implies the existence of an a with

$$au + (1-a)v > w.$$

C Algebra of combining

C:a $au + (1-a)v = (1-a)v + au$

C:b $a(bu + (1-b)v) + (1-a)v = cu + (1-c)v$, where $c = ab$.

Table II.2.2

Axiomatization of Utility

(Savage 1954)

In the following postulates (P1-P6), let S be the universal event, a set of states containing all the states of the world, with generic element s; let A and B be events, subsets of S; let f, f', g, g', denote consequences; let $\mathfrak{f}$, $\mathfrak{f}'$, $\mathfrak{g}$, $\mathfrak{g}'$ be acts, functions that attach a specific consequence f(s), f'(s), g(s), g'(s), respectively, to each state of the world s; and let $\mathfrak{f}_A$, $\mathfrak{f}_B$, $\mathfrak{g}_A$, $\mathfrak{g}_B$ denote acts conditioned on the event A, or B, obtaining.

P1 The relation <= is a simple ordering among acts.

P2 If $\mathfrak{f}$, $\mathfrak{g}$, and $\mathfrak{f}'$, $\mathfrak{g}'$ are such that:

1. in $\sim$B, $\mathfrak{f}$ agrees with $\mathfrak{g}$, and $\mathfrak{f}'$ agrees with $\mathfrak{g}'$,
2. in B, $\mathfrak{f}$ agrees with $\mathfrak{f}'$, and $\mathfrak{g}$ agrees with $\mathfrak{g}'$,
3. $\mathfrak{f} <= \mathfrak{g}$,

then $\mathfrak{f}' <= \mathfrak{g}'$

P3 If $\mathfrak{f}$ = g, $\mathfrak{f}'$ = g', and B is not null; then $\mathfrak{f} <= \mathfrak{f}'$ given B, if and only if g <= g'.

Table II.2.2

(continued)

P4 If f, f', g, g'; A, B; $\mathbf{f}_A$, $\mathbf{f}_B$, $\mathbf{g}_A$, $\mathbf{g}_B$ are such that:

1. f' <= f, g' <= g;

2a. $f_A(s)$ = f, $g_A(s)$ = g, for s in A

$f_A(s)$ = f', $g_A(s)$ = g', for s not in A;

2b. $f_B(s)$ = f, $g_B(s)$ = g, for s in B

$f_B(s)$ = f', $g_B(s)$ = g', for s not in B;

3. $\mathbf{f}_A$ <= $\mathbf{f}_B$;

then $\mathbf{g}_A$ <= $\mathbf{g}_B$.

P5 There is at least one pair of consequences f, f' such that f' < f.

P6 If $\mathbf{g} < \mathbf{h}$, and f is any consequence; then there exists a partition of S such that, if $\mathbf{g}$ or $\mathbf{h}$ is modified on any one element of the partition as to take the value f at every s there, the other values being undisturbed, then the modified $\mathbf{g}$ remains less than $\mathbf{h}$, or $\mathbf{g}$ remains less than the modified $\mathbf{h}$, as the case may require.

These two classes have been well researched and criticized (Friedman and Savage, 1948; Mosteller and Nogee, 1951; Coombs, 1975; Kahnemann and Tversky, 1979; Novick and Lindley, 1979; Novick, Turner, and Novick, 1981). The criticisms are discussed here in order to justify the selection of the fixed-state procedure of Novick and Lindley as the method for elicitation in this paper.

The two classes of procedures are based on the theory of expected utility. Both assume that there are three distinct outcomes or objects, which are ordered according to perceived value or utility. Two alternatives are presented: the object of middle value may be selected for sure, or a gamble may be selected involving the other two objects. If the for-sure alternative were selected, then the object of middle value would be obtained unconditionally, in this hypothetical choice. If the gamble were selected, the object of higher utility would be obtained with probability p and the object with lower utility would be obtained with probability 1-p. For certain objects and probabilities, the expected utilities of the two alternatives would be equal.

In the fixed-probability procedures, the value of the probability p is fixed, and the value of one of the three objects is varied until the utility of the for-sure alternative is equal to the expected utility of the gamble. In some cases it is the object of middle value, the for-sure object, that is varied; in some cases it is the higher-valued object that is varied. The middle-valued object is called the "certainty equivalent" of the gamble.

In the fixed-state procedures, the three object are fixed and the value of the probability that makes the expected utilities of the for-sure alternative and the gamble alternative equal is ascertained. This probability is called the "indifference probability," as one would be indifferent to a choice between the for-sure alternative and the gamble with such a probability.

The two classes of procedures may be compared on several levels. One is concerned with coherence in judgment (Novick and Lindley, 1979). If we elicit subjective judgments on related quantities, we want them to agree. In eliciting judgments in the assessment of utilities, we want to elicit more judgments than are necessary to estimate the unknown utility parameters. Thus, the assessee has an opportunity to see the effects of incoherent judgment.

The ease of coherence checking is not equal between the fixed-state and the fixed-probability procedures (Novick and Lindley, 1979). For example, in the fixed-probability procedure, fixing the probability at one half, we might find the certainty equivalent of an even-odds gamble between the worst state, say θ_0, and the best state, say θ_1. Calling this state $\theta_{.5}$ and assuming the utility of θ_0 to be 0 and the utility of θ_1 to be 1 (i.e., $u(\theta_0)=0$ and $u(\theta_1)=1$), we would set $u(\theta_{.5})=.5$. We could then elicit the certainty equivalents for even-odds gambles between θ_0 and $\theta_{.5}$ and between $\theta_{.5}$ and θ_1. As these certainty equivalents would have utilities .25 and .75, respectively, we might call them $\theta_{.25}$ and $\theta_{.75}$. If we were to then elicit the certainty equivalent for an even-odds gamble between $\theta_{.25}$ and $\theta_{.75}$, it would be rather obvious that it should be $\theta_{.5}$. Hence, we would not be getting observations that were independent enough to check for coherence.

The fixed-state procedures seem to not be affected as much by this difficulty in coherence checking. In the next section, several algorithms that have been developed for this purpose are discussed.

One criticism that appears to affect procedures from both classes stems from an "anchoring and adjustment" phenomenon (Tversky and Kahnemann, 1974). It is hypothesized that the first response acts as an "anchor" and subsequent responses are adjustments from the initial response. In addition, it appears that in both the fixed-state and the fixed-probability procedures, estimates of the extremes tend to be conservative. More research is being conducted in this area (Novick, Turner, and Novick, 1981).

Section 4. The Novick and Lindley Procedure

The procedure advocated by Novick and Lindley (1979) is a fixed-state procedure. In this section we present the procedure in detail. The implementation of the procedure (Isaacs and Novick, 1978) and its subsequent enhancement (Novick, Hamer, Libby, Chen, and Woodworth, 1980) are also discussed. In the following section, additional enhancements are suggested. Although much of the discussion of the implementation of the procedure is concerned with modelling and estimation, rather than elicitation, it is included here for reasons of continuity and completeness.

The Novick and Lindley fixed-state utility assessment procedure is based on the theory of expected utility in the following way. It assumes that we have a set _ of n+1 states in which we are interested; we will name these states $\theta_0, \theta_1, \ldots, \theta_n$. It assumes that we can order these states such that $\theta_0 <_p \theta_1 <_p \ldots <_p \theta_n$, where "<p" is interpreted as "less preferred than." Note that the preference ordering is assumed to be strict; it is not permitted that two states are equally preferred, or that one is indifferent in a choice between two states. (This assumption is for

explanatory convenience. Any implementation of the procedure should be able to determine when two states are preferentially equivalent and to adjust the model accordingly. This topic is discussed in more detail in Chapter VII when the procedure being advocated in this paper is presented.)

We will have occasion to discuss various states selected from the set $\underline{\theta}$. We will use θ_i, θ_j, and θ_k to denote any three states of the set such that $\theta_i <p\ \theta_j <p\ \theta_k$. We will use θ_{j+1} to mean the state that is preferred to θ_j but is preferred to no other state that is preferred over θ_j. Similarly, we use θ_{j-1} to denote the state to which θ_j is preferred but is preferred over all other states over which θ_j is also preferred.

The Novick-Lindley procedure presents a sequence of choices involving three states, θ_i, θ_j, and θ_k. In each of these hypothetical situations, the decision maker is given a choice between obtaining state θ_j for sure or participating in a gamble where θ_k would be obtained with probability p and θ_i would be obtained with probability 1-p. The purpose of the procedure is to determine a value for p so that the choice is indifferent to the decision maker. It is then assumed that the expected utilities of the two alternatives are equal. In other words, the equation

(II.4.1) $u(\theta_j) = p\ u(\theta_k) + (1-p)\ u(\theta_i)$

is true, where u(.) is the utility transformation. Since the indifference probability p varies depending on the situation, i.e., on the choice of i, j, and k, we will refer to a specific value as p_{ijk}.

It is this utility transformation that we are attempting to estimate. There are, however, some assumptions that we make about its form. First, it is a discrete transformation. We are not trying to find a parametric fit for the set of all possible states (our set $\underline{\Theta}$ is a subset of this set). We are merely attempting to find a value to associate with each element θ_i in the set $\underline{\Theta}$. Second, we associate a utility of 0 with the least preferred state in $\underline{\Theta}$ and a utility of 1 with the most preferred state. If we are aiming for an interval level of measurement with our utility transformation, this second assumption only restricts us enough to make the necessary estimates, because an interval level of measurement is unique only up to a linear transformation.

Since equation (II.4.1) involves three utilities and the second assumption specifies two utilities in the set to have fixed values, we can solve the system of equations in the form of equation (II.4.1) if we can arrange each equation so that it involves two utilities that we know and only one that we do not know. To accomplish this, it is convenient to select the states θ_i, θ_j, and θ_k so that they are 'adjacent,' i.e., we select θ_{j-1}, θ_j, and θ_{j+1} . Equation (II.4.1) then becomes

$$\text{(II.4.2)} \quad u(\theta_j) = p\, u(\theta_{j+1}) + (1-p)\, u(\theta_{j-1}).$$

Since the indifference probability p here depends on the selection of j only, we will refer to a specific value as p_j. We can now rearrange equation (II.4.2) so that it involves differences between the utilities of adjacent states

$$\text{(II.4.3)} \quad p_j\, u(\theta_j) + (1-p_j)\, u(\theta_j) = p_j\, u(\theta_{j+1}) + (1-p_j)\, u(\theta_{j-1})$$

$$\text{(II.4.4)} \quad p_j[u(\theta_{j+1}) - u(\theta_j)] = (1-p_j)\, [u(\theta_j) - u(\theta_{j-1})]$$

$$\text{(II.4.5)} \quad u(\theta_{j+1}) - u(\theta_j) = \left(\frac{1-p_j}{p_j}\right)[u(\theta_j) - u(\theta_{j-1})]$$

If we define $f_j = (1-p_j)/p_j$, then equation (II.4.5) becomes

$$\text{(II.4.6)} \quad u(\theta_{j+1}) - u(\theta_j) = f_j[u(\theta_j) - u(\theta_{j-1})].$$

Substituting equation (II.4.6) for j=j'-1 into the equation for j=j' and repeating the substitution, we arrive at

$$\text{(II.4.7)} \quad u(\theta_{j+1}) - u(\theta_j) = \left(\prod_{i=1}^{j} f_i\right)\, u(\theta_1)$$

since we have constrained $u(\theta_0) = 0$. If we then sum equations (II.4.7) over j from 0 to k (and defining $f_0 = 1$), we get

$$u(\theta_{k+1}) = \left[\sum_{j=0}^{k}\left(\prod_{i=0}^{j} f_i\right)\right] u(\theta_1). \tag{II.4.8}$$

If k+1=n then

$$u(\theta_n) = \left[\sum_{j=0}^{n-1}\left(\prod_{i=0}^{j} f_i\right)\right] u(\theta_1). \tag{II.4.9}$$

Since we have constrained $u(\theta_n) = 1$, we can calculate $u(\theta_1)$ from known quantities. Knowing $u(\theta_1)$, we can calculate the utility of the other states successively using equation (II.4.5).

This development does not use the concept of coherence checking. We are assuming only as much information as is needed to uniquely determine the utilities. To check coherence, we need to collect more information on the relationships among the utilities, i.e., we need to elicit indifference probabilities for more gambles. There are three methods that have been developed to check for coherence: local-coherence assessment, regional-coherence assessment, and least-squares assessment (Isaacs and Novick, 1978; Novick and Lindley, 1979; Novick, DeKeyrel, and Chuang, 1979; Novick, Hamer, Libby, Chen, and Woodworth, 1980; Novick, Turner, and Novick, 1981).

In the regional-coherence assessment procedure, adjustment is made for pairs of situations. Indifference probabilities are elicited for two situations in each of which a hypothetical choice between receiving a given state for sure and participating in a gamble is offered. Only four states are used for each pair of situations; the two situations have two states in common. The indifference probabilities for these two situations imply fixed indifference probabilities for the other two possible situations involving three out of the same four states (there are four unique combinations of three states chosen from four states when order is not a consideration). The indifference probabilities for all four situations are adjusted until they are acceptable; then another pair of situations is presented. When enough information has been collected to uniquely determine the utilities, the procedure stops.

In the local-coherence assessment procedure, adjustment is made for each situation in which a choice is made between a for-sure state and a gamble involving two other states. An indifference probability is elicited. From this information, two gambles are constructed, each involving two of the three states, such that a choice between the two gambles should be indifferent. If one of the gambles is preferred, the probability of one of the gambles is modified, thus changing the indifference probability in the original choice, so that neither gamble is preferred over the other. Once all the probabilities are acceptable, another set of three states is presented in the same way. When enough sets have been presented so that all the utilities can be uniquely determined, the iteration stops.

The least-squares assessment procedure allows adjustment only after all situations have been presented. For this reason, the procedure might be called a global-coherence assessment procedure as well. Indifference probabilities are elicited for each situation presented, as in the other assessment procedures, but more situations are presented than are necessary to uniquely determine the utilities. Therefore, it is unlikely that there is a solution that exactly fits all of the equations in expected utility. To estimate the utilities, a least-squares solution is calculated.

In the least-squares methodology, we assume that our observations, in this case the indifference probabilities p_j, contain some (unknown) amount of error, which may be random or systematic. Thus, the estimated model is not expected to fit the observations exactly. The parameters in the model are estimated to minimize the squared deviations of the observations from the corresponding values projected by the model.

Because of the importance to the estimation of the model, it is necessary to consider carefully the metric in which to make the comparison between the observations and the estimated model. One possible metric is that of the indifference probabilities themselves. Equation (II.4.1) can be rearranged to give

$$\text{(II.4.10)} \quad p_{ijk} = \frac{u(\theta_j) - u(\theta_i)}{u(\theta_k) - u(\theta_i)}$$

from which we can obtain estimates of the indifference probabilities conditional on the estimates of the utilities

$$\text{(II.4.11)} \quad \hat{p}_{ijk} = \frac{\hat{u}(\theta_j) - \hat{u}(\theta_i)}{\hat{u}(\theta_k) - \hat{u}(\theta_i)}$$

One could, then, estimate the utilities to minimize the sum of squared differences between p_{ijk} and $\hat{p}_{ijk}$.

Novick and Lindley (1979) have argued that the precision of the logodds of the indifference probabilities is the crucial factor in the analysis. Note that equation (II.4.1) can also be rearranged to produce

$$\text{(II.4.12)} \quad \frac{p_{ijk}}{1-p_{ijk}} = \frac{u(\theta_j) - u(\theta_i)}{u(\theta_k) - u(\theta_j)}$$

and the corresponding estimates

$$\text{(II.4.13)} \qquad \frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}} = \frac{u(\theta_j) - u(\theta_i)}{u(\theta_k) - u(\theta_j)} \quad .$$

The least-squares solution advocated by Novick and Lindley is, thus, the set of utilities that minimizes the equation

$$\text{(II.4.14)} \qquad \sum_{i,j,k} \left[\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) - \log\left(\frac{u(\theta_j) - u(\theta_i)}{u(\theta_k) - u(\theta_j)}\right) \right]^2 \quad .$$

Research has shown (Mayekawa, 1981) that use of the logodds metric does not appreciably improve the estimation over the use of the metric of the indifference probabilities or the arcsine metric.

Section 5. Enhancements to the Procedure

There are two enhancements noted here for the original implementation of the Novick and Lindley least-squares fixed-state utility assessment procedure (Isaacs and Novick, 1978). Although they are more relevant to the implementation of the procedure being advocated in this paper, they are mentioned here for continuity.

The first enhancement is concerned with the numerical method used in the nonlinear estimation. The method is a Newton-Raphson method (Dahlquist, Björk, and Anderson, 1974; Kennedy and Gentle, 1980) such that

(II.5.1) $\underline{u} - \underline{u}' = H\underline{g}$

where $\underline{u}$ is a previous estimate of the utilities, $\underline{u}'$ is the new estimate of the utilities, $\underline{g}$ is a vector of first derivatives of the function to be minimized (equation II.4.14), and H is the inverse of the matrix of second derivatives (the Hessian matrix) of the function to be minimized. All derivatives are with respect to the utility parameters. Table II.5.1 presents these first and second derivatives.

The original implementation of the Novick-Lindley procedure did not use the full expressions for the calculation of the Hessian matrix of second derivatives. In each of the sums for the second derivatives in Table II.5.1, only the first term within the brackets was used. The justification was based on the expectation that the difference between the two log terms would be sufficiently close to zero to make its contribution to the second term within the brackets negligible.

The first enhancement is to use the full expression for the second derivatives. It appears that this enhancement lessens the likelihood of obtaining utility estimates that are not monotonic. (The constraints imposed by the model theoretically guarantee that the utility estimates will be monotonic. With computational inaccuracies due to the

Table II.5.1

Least Squares Derivatives

$$F(\underline{u}) = \sum_{i<j<k}\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)^2, \quad Y_{ijk} = \log\left(\frac{p_{ijk}}{1-p_{ijk}}\right)$$

$$\frac{\partial}{\partial u_s}F(\underline{u}) = 2\sum_{i=s<j<k}\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_j-u_i}\right)$$

$$-2\sum_{i<j=s<k}\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_k-u_j}+\frac{1}{u_j-u_i}\right)$$

$$+2\sum_{i<j<k=s}\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_k-u_j}\right)$$

$$\frac{\partial^2}{\partial u_s \partial u_t}F(\underline{u}) = 2\sum_{i=s=t<j<k}\left[\left(\frac{1}{u_j-u_i}\right)^2 + \left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_j-u_i}\right)^2\right]$$

$$+2\sum_{i<j=s=t<k}\left[\left(\frac{1}{u_j-u_i}+\frac{1}{u_k-u_j}\right)^2\right.$$

$$\left.-\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\left(\frac{1}{u_k-u_j}\right)^2-\left(\frac{1}{u_j-u_i}\right)^2\right)\right]$$

$$+2\sum_{i<j<k=s=t}\left[\left(\frac{1}{u_k-u_j}\right)^2 - \left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_k-u_j}\right)^2\right]$$

$$-2\sum_{i=s<j=t<k}\left[\frac{1}{u_j-u_i}\left(\frac{1}{u_k-u_j}-\frac{1}{u_j-u_i}\right)\right.$$

$$\left.+\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_j-u_i}\right)^2\right]$$

$$+2\sum_{i=s<j<k=t}\left(\frac{1}{u_j-u_i}\right)\left(\frac{1}{u_k-u_j}\right)$$

$$-2\sum_{i<j=s<k=t}\left[\frac{1}{u_k-u_j}\left(\frac{1}{u_k-u_j}+\frac{1}{u_j-u_i}\right)\right.$$

$$\left.-\left(Y_{ijk}-\log\left(\frac{u_j-u_i}{u_k-u_j}\right)\right)\left(\frac{1}{u_k-u_j}\right)^2\right]$$

limited precision of the computing environment, however, the theoretical constraints are not sufficient, especially when the utilities of two states are very close.) This enhancement has been implemented in the 1980 release of the CADA Monitor (Novick, Hamer, Libby, Chen, and Woodworth, 1980). Other intricacies of the algorithm are discussed later.

The second enhancement is concerned with the algorithm used to calculate initial estimates for the utilities. Since the estimation algorithm is an iterative improvement algorithm, it requires a starting point. The current implementation of the Novick-Lindley procedure requires that all the adjacent gambles be used to calculate the initial estimates. Although it may be that the adjacent gambles are easiest and most accurate to elicit (Novick, Turner, and Novick, 1981) and therefore would always be available, this restriction is unnecessary. All that is required is a set of non-linearly-related gambles involving all the states. While the numerical analysis needed to check the conditions is more complicated than the currently implemented scheme, the increase in flexibility for the decision maker more than offsets this disadvantage. (The initial utility estimates are calculated in the metric of the indifference probabilities.) This enhancement has not been implemented in the CADA Monitor.

Chapter III. MULTIDIMENSIONAL SCALING

Section 1. Introduction

In this chapter, we will present the background material needed for the development of one approach to the utility assessment procedure that is the goal of this paper. To this purpose, the chapter is divided into four sections, including this one.

Section 2 presents a brief history of multidimensional scaling. The works of Hotelling (1933), concerning the principal components decomposition, of Eckart and G. Young (1936), concerning the approximation of a matrix by another of lower rank, and of G. Young and Householder (1938), concerning the representation of a set of points in terms of their mutual distances, stand as the foundation upon which accomplishments in the field of multidimensional scaling are built. Torgerson (1952) developed these ideas into a theory and method of multidimensional scaling, noting that the dimensions in psychological, as opposed to psychophysical, scaling are often unknown or confounded.

Kruskal and Shepard (Kruskal, 1964; Shepard and Kruskal, 1964; Shepard, 1966) expanded the concepts developed by Torgerson to include nonmetric stimuli (i.e., ordinal) as well as metric stimuli. The analysis of group scaling was developed by Bloxom (1968), Carroll and Chang (1970), and Schonemann (1972). This has been called the "individual differences" model. In this paradigm, it is assumed that data have been collected from several sources (e.g., individuals) that have a common underlying structure but different perceptions of that structure. The differences in perception may include heterogeneous weighting of the component axes or heterogeneous rotations of the axes.

Much of the current work in multidimensional scaling is being done by F. W. Young, et alia (de Leeuw and Pruzansky, 1976; de Leeuw, Young, and Takane, 1976; Takane, Young, and de Leeuw, 1977; Young, 1975; Young, de Leeuw, and Takane, 1976a; Young, de Leeuw, and Takane, 1976b; Young, Takane, and de Leeuw, 1978). A comprehensive theory of data and modelling has been developed that includes both metric and nonmetric scaling, and the many common models in multidimensional scaling such as the additive model, the regression model, the individual differences model, and the principal components model.

Section 3 presents the theory of data and modelling of F. W. Young, et alia. The theory of data is structured using the level of measurement (nominal, crdinal, interval, and ratio) and the process of measurement (discrete and continuous). This theory of data forms the basis for the scaling phase of the algorithm presented by F. W. Young, et alia, called Alternating Least Squares with Optimal Scaling (ALSOS). As the multiattribute utility assessment procedure presented in this paper uses this algorithm as a base structure, the ALSOS algorithm is discussed in detail.

In Section 4, various topics in multiattribute utility assessment that use the language and structure of multidimensional scaling are presented. Mcst of these applications use the additive model (e.g., Green and Wind, 1973; Kruskal, 1965). Some use of the individual differences model has also been made (Green and Wind, 1973).

Section 2. The History of Multidimensional Scaling

One of the initial attempts to view a matrix as a projection from some set of axes different from the apparent axes of the observed measurements is due to Hotelling (1933). In his paper, Hotelling shows that one can construct a set of axes such that each axis represents the dimension of largest variation in the observed values that is orthogonal to the other axes.

Hotelling's paper, and results due to Courant and Hilbert (Eckart and G. Young, 1936), were used by Eckart and G. Young (1936) to determine the closest approximation to a given matrix so that the approximation is of lower rank. The solution is based on two results by Courant and Hilbert. The first states that any real matrix may be decomposed:

> For any real matrix A, two orthogonal matrices V and U can be found so that L = VAU' is a real diagonal matrix with no negative elements.

This theorem shows that a decomposition of a real matrix exists. The second theorem, also due to Courant and Hilbert, shows the conditions under which a matrix of lower rank that closely approximates a given matrix may be found:

> If AB' and B'A are both symmetric matrices, then and only then can two orthogonal matrices V and U be found such that L = VAU' and M = VBU' are both real diagonal matrices.

The solution to the problem of finding an approximation to a matrix that has a lower rank, say r, is to decompose the matrix as indicated in the first theorem and use the first r rows and columns of the resulting decomposition and the second theorem to construct the approximating matrix.

Another basic result is the representation of a set of points in Euclidean space in terms of their mutual distances (G. Young and Householder, 1938). In this paper, it is shown that under certain conditions, a set of numbers may be conceived as mutual distances of a set of points in Euclidean space. Matrices are also found in the paper whose ranks determine the smallest Euclidean space containing such points. The results of Eckart and G. Young are then applied so that the representation of the set of numbers may be approximated by a set of points in an Euclidean space of lower dimensionality.

These results may be applied to any set of numbers that may be interpreted as measurements of distances among a set of objects. The implications of these results were expressed by Torgerson (1952) as follows:

> The traditional methods of psychophysical scaling presupposes knowledge of the dimensions of the area being investigated. The methods require judgments along a particular defined dimension, i.e., A is brighter, twice as loud, more conservative, or heavier than B. The observer, of course, must know what the experimenter means by brightness, loudness, etc. In many stimulus domains, however, the dimensions themselves, or even the number of relevant dimensions, are not known. What might appear intuitively to be a single dimension may in fact be a complex of several. Some of the intuitively given dimensions may not be necessary -- it may be that they can be accounted for by linear combinations of others. Other dimensions of importance may be completely overlooked. In such areas the traditional approach is inadequate.

The results of G. Young and Householder are not directly applicable to the type of problems ennunciated by Torgerson. The work of the former authors presupposes infallible data.

The data from the oft-cited real world that Torgerson wishes to analyze are likely to be distorted by errors of measurement.

The crucial aspect is in the measurement of the distances between the objects. The origin of the Euclidean space being determined is unknown. Young and Householder do not address this issue, instead choosing one of the objects to represent the origin. There are, then, an infinity of solutions, but each is a Euclidean transformation of any solution obtained by arbitrarily selecting an origin.

When the data are fallible, this solution is no longer reliable. Selecting different objects to represent the origin, each with its unique error of measurement, may lead to qualitatively, as well as quantitatively, different representations in Euclidean space. In particular, they are not likely to be linear transformations of each other. A solution to this problem, posed by Torgerson, is to choose the centroid of the objects as the origin of the Euclidean space.

Torgerson's work assumes, however, that the operations of arithmetic may be validly applied to the data measured. For this reason, the procedure is called a 'metric' algorithm. The field of multidimensional scaling was expanded to include nonmetric procedures, in particular those that rely on ordinal judgments about the stimuli, by the use of a scaling transformation. The scaling transformation is estimated as a monotone function of the judgements, so to preserve the ordinal characteristics of the data. The crucial aspect of the transformation is the estimation of a true zero on the scale of measurement so that the operations of arithmetic are valid. With the judgments converted to distance-like, or scalar-product-like, measures, the results of Eckart and G. Young, and G. Young and Householder, may be applied, and a Euclidean space of low dimensionality may be constructed that adequately represents the objects of interest.

This extension of multidimensional scaling was pioneered by Kruskal and Shepard (Kruskal, 1964; Shepard and Kruskal, 1964; Shepard, 1966). An additional important contribution to this area is due to Guttman (Guttman, 1968).

All of the above works presuppose a single, unified point of view. If data are collected from several sources, the individual information is lost, as the emphasis is on some type of composite. Attempts to isolate an underlying, common point of view yet preserve some of the individual information led to the development of the "individual differences" model of multidimensional scaling (Tucker and Messick, 1963; Bloxom, 1968; Horan, 1969; Carroll and Chang, 1970; Schonemann, 1972; Takane, F. W. Young, and de Leeuw, 1977).

As in the previous models, a set of points in a Euclidean space of some dimension is determined using a composite point of view. In addition, a set of weighting factors are estimated for each individual. These weighting factors alter the unit of measurement along each of the dimensions of the common Euclidean space according to the importance of that dimension as perceived by the individual. Some of the individual differences models allow reflections of the dimensions or rotations of the axes.

An algebraic solution to the individual differences model has been proposed (Schonemann, 1972). Although the algebraic solution is of some theoretical significance, its practical application is limited, again because of the instability problems of fallible data. In the above-mentioned paper, Schonemann states:

> We do not necessarily recommend its use in preference over presently available iterative algorithms. Algebraic solutions sometimes have a tendency to become unstable in the fallible case, and it is therefore often safer to replace them in actual applications by algorithms which have well understood optimality properties...

The approach in this paper follows the above suggestion in that an iterative solution, not an algebraic one, is proposed.

A unified theory of data has been proposed for multidimensional scaling by F. W. Young, et alia (Young, 1975; de Leeuw, Young, and Takane, 1976; Young, de Leeuw, and Takane, 1976a; Young, de Leeuw, and Takane, 1976b; Takane, Young, and de Leeuw, 1977; Young, Takane, and de Leeuw, 1978). The development of multidimensional scaling models and algorithms had been unfocused, but the theory of data developed by F. W. Young, et alia, sets a structure in which the various models fit. The next section presents this theory of data in detail.

Section 3. The ALSOS Algorithm

Young, et alia (F. W. Young, de Leeuw, and Takane, 1976a), have defined optimal scaling as follows:

> Optimal scaling is a data analysis technique which assigns numerical values to observation categories in a way which maximizes the relation between the observations and the data analysis model while respecting the measurement character of the data.

Scaling techniques other than that proposed by F. W. Young, et alia, have been used (e.g., Torgerson, 1952; Kruskal, 1964; Guttman, 1968) which conform to this definition, but the Alternating Least Squares with Optimal Scaling (ALSOS) is the most coherently stated and supported by a theory of data.

The structure of the ALSOS algorithm consists of a data space, an optimal scaling space, a model space, and a parameter space. The data space and the optimal scaling space are related by an optimal scaling transformation. The optimal scaling space can be viewed as a projection of the data space onto a numerical space restricted by the measurement characteristics of the data. The model space is a least squares projection of the optimal scaling space, subject to the restrictions imposed by the model. The model space and the parameter space are related by a combination rule, or formula, which defines the model.

The alternating least squares algorithm consists of dividing all the parameters into two mutually exclusive and exhaustive sets. One set is then taken to be fixed, and a solution is calculated for the second set. This solution is then taken as fixed for the second set, and a solution is calculated for the first set. Hence the term "conditional least squares" is used for this algorithm (de Leeuw, Young, and Takane, 1976), as the least squares solution at each step is conditional on fixed (possibly unstable) values of the other parameters. This algorithm has also been called "block relaxation" (Céa and Glowinski, 1973; de Leeuw, F. W. Young, and Takane, 1976). The process is iterated until convergence criteria are satisfied.

In the optimal scaling phase of the algorithm, the observations are viewed as categorical and each observation category is represented by a parameter. The number and the nature of the categories, and thus the parameters, are determined by the measurement characteristics of the data. These are discussed more fully later in this section.

In the model estimation phase of the algorithm, the parameters of the model are divided into mutually exclusive and exhaustive sets as well. A least-squares solution is calculated for each set in turn, conditional on fixed values for the other sets of parameters. The overall algorithm alternates between the optimal scaling phase and the model estimation phase until convergence criteria are met.

Convergence of the ALSOS algorithm has been demonstrated (de Leeuw, F. W. Young, and Takane, 1976; de Leeuw, undated). It must be noted, though, that convergence is guaranteed only to a solution, not necessarily to the globally optimal solution. The developers of the algorithm, however, appear to be satisfied that the globally optimal solution is nearly always obtained (de Leeuw, F. W. Young, and Takane, 1976).

The applicability of the algorithm to a wide range of models is indicated by the separation of the optimal scaling phase and the model estimation phase. In fact, F. W.

Young, et alia (F. W. Young, de Leeuw, and Takane, 1976a), have stated:

> If a least squares method is known for analyzing quantitative data then a least squares method can be constructed for analyzing qualitative data.

This, then, is the basis for applying multidimensional scaling methods to the problem of assessing utilities in multiple dimensions. In order to support this application, a closer look at the optimal scaling transformations and the measurement characteristics that imply them is needed.

The optimal scaling transformation is governed by the measurement characteristics of the data. These measurement characteristics have been organized into a theory of data by F. W. Young, et alia (F. W. Young, de Leeuw, and Takane, 1976a,b), described in terms of level of measurement and process of measurement. The levels of measurement, as described previously, are binary, nominal, ordinal, interval, and ratio. The latter two levels are often combined as numerical data. The processes of measurement are discrete and continuous.

As mentioned above, all observations are interpreted as categorical. This is justified in the sense that the procedure by which observations are obtained is limited by the finite precision of measurement and recording. The level of measurement, then, is concerned with the relationships of the observations among the categories, whereas the process of measurement is concerned with the relationships within the categories.

The binary level of measurement postulates two categories, or more strictly, that an observation either has certain characteristics or it does not. In the nominal level of measurement, there can be several categories, but the only restriction governing the relationships among the categories is that the categories are not equivalent. In the ordinal level of measurement, the categories are assumed to be ordered in some way, but there are no restrictions on the distances between pairs of observations. In the interval level of measurement, the optimally-scaled value is assumed to be a polynomial function of the observation value, including an optimally-scaled origin. In the ratio level of measurement the optimally-scaled value is also assumed to be a polynomial function of the observation value, but the origin of the optimal scaling space is assumed to be the same as that for the observation space.

For the discrete process of measurement, the optimal scaling space for all observations within a category is represented by one number. On the other hand, in the continuous process of measurement, the optimal scaling space within a category is represented by an interval. Thus, for the discrete case, observations in the data space that are measured to be equal, or that are equally preferable, are transformed so that the optimally-scaled values are equal (i.e., ties must remain tied). In the continuous case, observations that are measured to be equal need not have equal optimally-scaled values; however, the optimally-scaled values of all observations within a category will lie in some interval. In both cases, observations that are measured not equal may, or may not, have optimally-scaled values that are equal, or not equal.

F. W. Young, et alia, have related this theory of data to other works. The discrete-nominal case had been previously developed by Fisher (1946). The continuous-nominal case is handled by a two-phase method, the first a discrete-nominal solution and the second a continuous-ordinal solution. The ordinal cases are similar to Kruskal's two approaches (Kruskal, 1964). The discrete-ordinal case is like Kruskal's secondary approach, whereas the continuous-ordinal case is like his primary approach.

The discrete-interval transformation is a polynomial linear regression (linear in the parameters). The optimally-scaled values are polynomial functions of the data. In the continuous case, the polynomial regression is followed by an estimation of the interval boundaries. The ratio cases are the same as the interval cases, with the exception that the regression is through the origin.

Modelling of the scaled data is equally important in Young's theory of data. Various models are discussed in the next section in exposition of multidimensional scaling techniques in multiattribute utility assessment and in the next chapter in detail.

Section 4. Some Examples

Some of the models of multidimensional scaling have been used for the purpose of estimating utilities. Some, such as the additive model, have straight-forward counterparts in the utility assessment literature. These will be discussed more fully in the next chapter on conjoint measurement. Others, such as the individual differences model, do not have direct analogies in the utility assessment field, but do correspond to generalizations of models in conjoint measurement. These, too, will be discussed more fully in the next chapter. The purpose of this section is to reference some examples of the use of these models.

The additive model is the simplest of the models. Its thesis is that the overall utility structure is composed solely of the sum of the utilities along the dimensions of the data space. In mathematical terms, the model may be stated as follows:

(III.4.1) $u(x_1, x_2, \ldots, x_n) = u_1(x_1) + u_2(x_2) + \ldots + u_n(x_n)$

where n is the number of dimensions, x_i is a value along the ith dimension, and $u_i(x_i)$ is the utility at value x_i along the ith dimension.

This model has been used extensively in the literature (e.g., Kruskal, 1965; Green and Wind, 1973; de Leeuw, F. W. Young, and Takane, 1976; Keeney and Raiffa, 1976; Humphreys and Wisudha, 1979). The model is analogous to the main-effects analysis of variance, assuming that the interaction terms are all zero. The difficulty with the model is its assum-ʽions about additivity. It presupposes that the utility structures along parallel lines through the data space (e.g., parallel to the axis of one dimension) are simply translations of each other. This is a strong assumption to make and it should be tested carefully.

A second model commonly purported is the individual differences model (e.g., Tucker and Messick, 1963; Bloxom, 1968; Horan, 1969; Carroll and Chang, 1970; Schonemann, 1972; Green and Wind, 1973; Takane, F. W. Young, and de Leeuw, 1977). In this model, we are presupposing that the data are collected from several, not necessarily comparable sources and we are estimating a Euclidean subspace which represents a common point of view. We also estimate individual weights for the dimensions of the Euclidean space which represent the importances of the dimensions as perceived by the individuals. As a mathematical model, it can be represented as follows:

(III.4.2) $D_i = X'W_iX$

where D_i is a matrix of scaled data for individual i, X is a matrix of coordinates in some Euclidean space for the set of stimuli, and W_i is a matrix of weights for individual i. In some models, W_i is constrained to be diagonal so that the model allows only differences of perception along the axes of the common space represented by the matrix X. In other formulations, the matrix W_i may be a general orthogonal matrix so that the individual differences may be manifested as rotations of the set of axes of the common space.

For examples of the use of individual differences models in utility assessment situations, see Green and Wind (1973).

Chapter IV. CONJOINT MEASUREMENT

Section 1. Introduction

The field of conjoint measurement is concerned with the foundations of measurement. Assuming only the existence of sets of objects and basic relationships among them, the existence and uniqueness of numerical scales may be established. Several axiomatizations have been presented in the literature that result in different structures involving multiple dimensions.

Conjoint measurement is closely related both to multidimensional scaling and to utility estimation. It has been claimed that multidimensional scaling is a special case of conjoint measurement (Tversky, 1967a; F. W. Young, 1972). It can also be seen that utility estimation, in various formulations, is a special case of conjoint measurement (Tversky, 1967a; F. W. Young, 1972; Krantz, Luce, Suppes, and Tversky, 1972). One of the goals of this paper is to present utility estimation in a general formulation of conjoint measurement.

The field of conjoint measurement is also relevant to this paper for its model formulations. The two most widely used models for multiple-attribute utility functions, the additive and the multiplicative models, have been extensively investigated in the literature (Krantz, Luce, Suppes, and Tversky, 1972; Keeney and Raiffa, 1976). The additive model presupposes that a multiple-attribute utility function is a (possibly weighted) sum of the single-attribute utility functions of its components. The multiplicative model presupposes that a multiple-attribute utility function is a product of the single-attribute utility functions of its components.

In Section 2, the additive and multiplicative models are discussed. An axiomatization of measurement for each model is briefly presented. It must be realized that other axiomatizations leading to the same model are possible, and that the axiomatizations selected are for illustrative purposes and not particularly for comparison. Some implementations of these two models are also referenced and the disadvantages of the models are outlined.

In Section 3, two more general theories of conjoint measurement are discussed (Tversky, 1967a; F. W. Young, 1972). Both are concerned primarily with polynomial conjoint measurement, i.e., the model of a multiple-attribute utility function is a general polynomial function of the utility functions of its components. The second theory, unlike the first, is not an axiomatization, although it is more general in that it allows more general functions for models than polynomials. The additive and the multiplicative models are special cases of these theories. The utility assessment procedure of Novick and Lindley (1979) is also related to the formulation of Young, thus establishing the connection with a concrete example and providing the motivation for the development that follows in the next chapter.

Section 2. Additive and Multiplicative Models

In this section, two axiomatizations of conjoint measurement are presented. The first is an axiomatization of additive conjoint measurement (Luce and Tukey, 1964). The second is an axiomatization of multiplicative conjoint measurement (Roskies, 1965).

In the additive axiomatization, the resulting representation of the measurement scale of a multiple-attribute consequence set is a sum of measurement scales of the individual attributes. For example, let A, P, ..., K be disjoint sets of consequences; then, if the consequence space of interest is a subset of the product space AxBx...xK,

(IV.2.1) $u(a,b,\ldots,k) = u_A(a) + u_B(b) + \ldots + u_K(k)$.

The axiomatizations leading to this representation are presented in Table IV.2.1. This is the original axiomatization; there have been several extensions that are not presented here (Krantz, 1964; Luce, 1966a; Krantz, Luce, Suppes, and Tversky, 1972). Some of these extensions provide for the k-dimensional representation illustrated above. Other extensions generalize the notation. The characterization is essentially the same.

The key axiom to the additive representation is Axiom 3, the Cancellation Axiom. It is best illustrated in the following three by three table :

	P	X	Q
A	(A,P)	(A,X)	(A,Q)
F	(F,P)	(F,X)	(F,Q)
B	(B,P)	(B,X)	(B,Q)

The axiom assumes that the orderings along the diagonals, from the upper-left-hand corner to the lower-right-hand corner, are consistent.

This is a strong assumption to make, and it leads to a strong representation. In the additive model, the contribu-

Table IV.2.1

The Luce-Tukey Axiomatization

(Luce and Tukey 1964)

Let $\mathcal{A}$ be a set with typical elements A, B, C, ..., F, G, H, ... and $\mathcal{P}$ a set with typical elements P, Q, R, ..., X, Y, Z; then $\mathcal{A} \times \mathcal{P}$ consists of pairs (A,P), (A,Q), (B,Q), etc. Let $\geq$ be a binary relation on such pairs. (Thus $\geq$ is equivalent to a subset of $(\mathcal{A}\times\mathcal{P})\times(\mathcal{A}\times\mathcal{P})$).

(VA) Ordering Axiom (Axiom 1). $\geq$ is a weak ordering, i.e.,

(VB, Reflexivity) (A,P) $\geq$ (A,P) holds for all A in $\mathcal{A}$ and P in $\mathcal{P}$;

(VC, Transitivity) (A,P) $\geq$ (B,Q) and (B,Q) $\geq$ (C,R) imply (A,P) $\geq$ (C,R);

(VD, Connectedness) Either (A,P) $\geq$ (B,Q) or (B,Q) $\geq$ (A,P), or both.

(VE) Definition. For A, B in $\mathcal{A}$ and P, Q in $\mathcal{P}$, (A,P) = (B,Q) if and only if (A,P) $\geq$ (B,Q) and (B,Q) $\geq$ (A,P); (A,P) > (B,Q) if and only if (A,P) $\geq$ (B,Q) and not((B,Q) $\geq$ (A,P)).

(VF) Solution (of Equations) Axiom (Axiom 2). For each A in $\mathcal{A}$ and P, Q in $\mathcal{P}$, the equation (F,P) = (A,Q) has a solution F in $\mathcal{A}$, and for each A, B in $\mathcal{A}$ and P in $\mathcal{P}$, the equation (A,X) = (B,P) has a solution X in $\mathcal{P}$.

Table IV.2.1

(continued)

(VG) Cancellation Axiom (Axiom 3). For A, F, B in $\mathcal{A}$ and P, X, Q in $\mathcal{P}$, $(A,X) \geq (F,Q)$ and $(F,P) \geq (B,X)$ imply $(A,P) \geq (B,Q)$.

(VIA) Characterization. A doubly infinite series of pairs $\{A_i,P_i\}$, $i=0,\pm 1, \pm 2, \ldots$, with A_i in $\mathcal{A}$ and P_i in $\mathcal{P}$, is a dual standard sequence provided that

(VIB) $(A_m,P_n) = (A_p,P_q)$ whenever $m+n=p+q$ for positive, zero, or negative integers m, n, p, and q. A dual standard sequence is trivial if for all i either $A_i = A_0$ or $P_i = P_0$, in which case both hold by transfer.

(VIC) Archimedean Axiom (Axiom 4). If $\{A_i,P_i\}$ is a non-trivial dual standard sequence, B is in $\mathcal{A}$ and Q is in $\mathcal{P}$, then there exist (positive or negative) integers n and m such that $(A_n,P_n) \geq (B,Q) \geq (A_m,P_m)$.

[illegible] the other attributes. Luce and Tukey have stated ([illegible]):

> More generally, a question raised throughout the social and behavioral sciences is whether two independent variables contribute independently to an overall effect or response. The usual approach is to attach to each pair of values of the variables a numerical measure of effect that preserves the order of effects and then to test for independence using an additive statistical model, probably one of the conventional analysis of variance models. When dependence (interaction) is shown to exist, one is uncertain whether the dependence is real or whether another measure would have shown a different result. Certain familiar transformations are often applied in an effort to reduce the danger of the second possibility, but they are unlikely to approach exhausting the infinite family of monotonic transformations, so that one cannot be too sure of the reality of an apparent interaction. Our results show that additive independence exists provided that our axioms are satisfied; of these, the most essential one from a substantive point of view is the cancellation axiom, which is also a necessary condition for an additive representation to exist. Thus, one could test the cancellation axiom by examining a sufficiently voluminous body of ordinal data directly, without introducing any numerical measures and, thereby, test the primary ingredient in additive independence. In some applications this should be more convincing than present techniques.

Comments about the adequacy of the additive representation in general situations are presented later in this section. Suffice it to mention here that the assumption of additivity is not one lightly accepted, and one needs alternatives for those situations when the additive representation is not appropriate because of the violation of one or more of the axioms.

Closely related to the additive representation is the multiplicative representation. The axioms of the Roskies system (Roskies, 1965) are listed in Table IV.2.2. There is a clear analogy between the Luce-Tukey system and the Roskies system. Axioms 1 through 3 of each system serve analogous purposes. Axiom 6 of the Roskies system is stated to be the same as Axiom 4 of the Luce-Tukey system. The major difference are the axioms in the Roskies system to establish the idea of zero.

The representation implied by the Roskies axiomatization is a multiplicative one, as follows:

$$\text{(IV.2.2)} \quad 1 + ku(A_1 \times A_2 \times \ldots \times A_k) = \prod_{i=1}^{k} \left[1 + k_i k u_i(A_i) \right] .$$

Note that this may be converted to an additive representation by the use of the log transformation and the realization that the scales u, u_1, ..., u_k are unique only up to a positive linear transformation.

These two representations, the additive model and the multiplicative model, are prominent in the field of conjoint measurement and in the field of utility assessment. Specific examples may be found in Fishburn (1966), Pollak (1967), Green and Wind (1973), Keeney and Sicherman (1975), Keeney and Raiffa (1976), and Humphreys and Wisudha (1979),

Table IV.2.2

The Roskies Axiomatization

(Roskies 1965)

Let $\mathcal{A}$ be a set with elements A, B, C, ..., F, G, ... and let $\mathcal{P}$ be a set with elements P, Q, R, ..., X, Y, ... $\mathcal{A} \times \mathcal{P}$ consists of pairs (A,P), (B,P), (C,X), etc.

Axiom 1. Ordering. Let $\geq$ be a binary relation on $\mathcal{A} \times \mathcal{P}$ such that $\geq$ is a weak ordering; that is, for all A, B in $\mathcal{A}$ and P, Q in $\mathcal{P}$,

(a) $(A,P) \geq (A,P)$;

(b) if $(A,P) \geq (B,Q)$ and $(B,Q) \geq (C,R)$ then $(A,P) \geq (C,R)$;

(c) either $(A,P) \geq (B,Q)$ or $(B,Q) \geq (A,P)$.

Definition. $(A,P) = (B,Q)$ if and only if $(A,P) \geq (B,Q)$ and $(B,Q) \geq (A,P)$; $(A,P) > (B,Q)$ if and only if not$((B,Q) \geq (A,P))$.

Definition. Since the ordering $\geq$ is weak, it partitions $\mathcal{A} \times \mathcal{P}$ into equivalence classes defined by =. From now on, we work with the equivalence classes, which we continue to denote by (A,P).

Table IV.2.2

(continued)

Definition of zero. Let

$\mathcal{A}_0 = \{A_0 \mid (A_0, P) = (A_0, Q)$ for all P, Q in $\mathcal{P}\}$,

$\mathcal{P}_0 = \{P_0 \mid (A, P_0) = (B, P_0)$ for all A, B in $\mathcal{A}\}$,

$\mathcal{J} = \{(A, P) \mid$ either A is in $\mathcal{A}_0$ or P is in $\mathcal{P}_0$ or both$\}$.

Axiom 2. Solution of equations. Let A, B be in $\mathcal{A}$ and P, Q be in $\mathcal{P}$. If B is not in $\mathcal{A}_0$, there exists X in $\mathcal{P}$ such that $(A, P) = (B, X)$. If P is not in $\mathcal{P}_0$, there exists F in $\mathcal{A}$ such that $(F, P) = (B, Q)$.

Axiom 3. Cancellation. Let A, A', A'' be in $\mathcal{A}$ and P, P', P'' be in $\mathcal{P}$. If A' is not in $\mathcal{A}_0$ or if P' is not in $\mathcal{P}_0$, and if $(A, P') = (A', P'')$ and $(A', P) = (A'', P')$, then $(A, P) = (A'', P'')$.

Axiom 4. Sign. Let P, Q be not in $\mathcal{P}_0$. If $(A,P)s(A,Q)$ for some A not in $\mathcal{A}_0$, then $(B,P)s(B,Q)$ for all B not in $\mathcal{A}_0$; similarly, if A, B be not in $\mathcal{A}_0$ and if $(A,P)s(B,P)$ for some P not in $\mathcal{P}_0$, then $(A,Q)s(B,Q)$ for all Q not in $\mathcal{P}_0$, where $(A,P)s(B,Q)$ if $(A,P) > 0$ and $(B,Q) > 0$ or if $(A,P) < 0$ and $(B,Q) < 0$. When $(A,P)s(A,Q)$, we write PsQ; s is an equivalence relation on $\mathcal{P} - \mathcal{P}_0$. We define AsB similarly, and we define $P{-}sQ$ and $A{-}sB$ in the obvious way: $(A,P){-}s(B,Q)$ means that $(A,P) > 0$ and $(B,Q) < 0$ or $(A,P) < 0$ and $(B,Q) > 0$.

Table IV.2.2

(continued)

Axiom 5. Let A, B be not in [illegible]$_a$, P, Q not in [illegible]$_p$.

If PsQ and $(A,P) \geq (B,P)$ then $(A,Q) \geq (B,Q)$;

if P-sQ and $(A,P) \geq (B,P)$ then $(A,Q) \leq (B,Q)$;

if AsB and $(A,P) \geq (A,Q)$ then $(B,P) \geq (B,Q)$;

if A-sB and $(A,P) \geq (A,Q)$ then $(B,P) \leq (B,Q)$.

Axiom 6. The Archimedean property as defined by Axiom 4 of the Luce-Tukey system holds on $(\mathcal{A}^+ \times \mathcal{P}^+)$.

among others.

Section 3. Polynomial Conjoint Measurement

In this section, two general theories of conjoint measurement are discussed. One is a polynomial representation axiomatization (Tversky, 1967a). The other (F. W. Young, 1972) is not an axiomatization, and is not restricted to polynomial representations, but is a presentation of a general theory of conjoint measurement models.

The axiomatization by Tversky is presented here for completeness of background material. The axiomatization is not discussed in detail, but is merely presented as theoretical support for the representation, or model, that is the subject of this paper. The paper by Young is presented to provide justification for the model developed in the next chapter.

The axiomatization by Tversky is a generalization of those by Luce and Tukey and by Koskies. As Tversky states:

> The present theory generalizes conjoint measurement in five major respects. (a) It is formatted in terms of partial rather than fully ordered data. (b) It applies to both ordinal and numerical data. (c) It is applicable to finite as well as infinite data structures. (d) It provides a necessary and sufficient condition for measurement. (e) This condition applies to any

polynomial measurement model; that is, any model where each data element is expressed a a specified real-valued, order-preserving polyncmial function of its components.

The first respect is important because it allows two data elements to be uncomparable. In previous axiomatizations, it was necessary that for every two data elements one be at least as preferable as the other. In the applications considered in this paper, utility assessment, it may be that two data elements may not be comparable because of the unlikelihood or the incomprehensibleness of one or the other of the situations that the data elements represent. This generalization allows us to ignore such a problem.

The second respect provides a theoretical justification for the levels-of-measurement data structuring used by F. W. Young, et alia, as presented in the previous chapter, and adapted in this paper. Although the proofs of the representation provided by Tversky are not directly applicable to all the data structures described by Young, in particular the nominal level of measurement, they do apply to most of the situations, and the general structure of the theory appears to fit in all cases.

The third respect is important because it generalizes the solvability axioms, which were crucial in the previous axiomatizations. In the Luce-Tukey and in the Roskies systems, it is necessary that the data structures be dense enough so that solutions to equivalence relations (=) exist in all cases. The formulations of the solvability axioms required that the data structure be infinite. The Tversky generalization relaxes this restriction, allowing greater flexibility in application.

The previous axiomatizations did not directly address the question of necessity and sufficiency, and therefore should be considered to represent sufficient conditions only. The Tversky theory does address this question.

The fifth respect allows less restrictive models than the previous representations. If x = (a,b,...,k) is a data element in AxBx...xK where a is in A, b is in B, ..., k is in K; l, m, ..., n are integers; then the representation can be displayed as

(IV.3.1) $$f(x) = \sum f_A^{l}(a) * f_B^{m}(b) * \ldots * f_K^{n}(k)$$

where the summation is over some subset of all combinations of l, m, ..., n. Formally, Tversky states:

> A data structure D = < [illegible], >o > is a system where is a subset of the product set AxBx...xK of some finite number of disjoint sets A, B, ..., K, and is partially ordered under >o. That is, >o is a binary relation defined on [illegible] which satisfies the following conditions for all x, y, z in [illegible]:
>
> (i) Reflexivity, x >o x;
>
> (ii) Transitivity, x >o y and y >o z imply x >o z;
>
> x =o y is defined as x >o y and y >o x; x >o y is defined as x >o y and not y >o x.

A data structure D is said to satisfy a polynomial measurement model M whenever there exists a real-valued function f defined on D and real-valued functions f_A, f_B, ..., f_K defined on factors A, B, ..., K such that, for any data element (a,b,...,k):

(i) $f(a,b,\ldots,k) = M(f_A(a), f_B(b), \ldots, f_K(k))$ where M is a polynomial function of its arguments, that is, a specified combination of sums, differences, and products of the functions f_A, f_B, ..., f_K;

(ii) for all $x = (a,b,\ldots,k)$, $x' = (a',b',\ldots,k')$, $x >_o x'$ implies $f(x) > f(x')$ and $x =_o x'$ implies $f(x) = f(x')$ where $>_o$ and $=_o$ denote the order observed in the data.

Tversky notes that the general polynomial conjoint measurement model encompasses a wide variety of well-known measurement models. Included as examples are Hull's and Spence's performance models as cited in Hilgard (1956), the Bradley-Terry-Luce choice model (Luce, 1959), the multidimensional scaling models, and Savage's subjective expected utility model (Savage, 1954).

Although not an axiomatization, and thus not as strongly placed on theoretical measurement foundations, is the generalization of conjoint measurement offered by Young (F. W. Young, 1972). Young presents a general functional form,

(IV.3.2) $h(x) = h_1(h_2(x_{i.}, x_{j.}))$, $i = 1,\ldots,I$, $j = 1,\ldots,J$

where the notation $x_{i.}$ is used to denote the ith row of X, and X is a matrix of coordinates used to spatially represent a set of objects.

Young then shows how various models presented in the literature can be represented by this general functional form. Included are the Euclidean and Minkowski distance scaling (Kruskal, 1964; Shepard, 1966; Guttman, 1968), multidimensional unfolding (Coombs, 1964) and a generalization to Minkowski metrics, monotone analysis of variance (Kruskal, 1965) which is related to the additive conjoint measurement model (Luce and Tukey, 1964), polynomial conjoint measurement (Tversky, 1967a), non-metric factor analysis (Shepard, 1966; Lingoes and Guttman, 1967; Kruskal and Shepard, 1974), subjective expected utility (Savage, 1954), and the Bradley-Terry-Luce choice model (Luce, 1959).

Using Tversky's theory, being applicable to numerical data as well as ordinal data, it is easy to see that Young's formulation applies equally well to metric factor analysis (Libby, 1979), principle components analysis (Young, Takane, and de Leeuw, 1978), regression analysis (Young, de Leeuw, and Takane, 1976b), analysis of variance (Kruskal, 1964; de Leeuw, Young, and Takane, 1976), and analysis of covariance structures (Libby, 1979).

We can also formulate the model of utilities used by Novick and Lindley (1979) using Young's theory, in the same way that Tversky and Young formulate similar models. Although not strictly a polynomial model, the matrix of indifference probabilities can be likened to Young's matrix of similarities which is fitted by a function, in this case the ratio of certain differences, of the utilities.

There are two relevant conclusions in the paper by Tversky. The first is

> The generality of the present theory, however, stems not only from the fact that most measurement models proposed can be represented as polynomial functions but also from the well-known result that any continuous real-valued function on a closed bounded region can be uniformly approximated arbitrarily closely by a polynomial function.

This certainly applies to the applications of interest for this paper, namely utility assessment. The second conclusion is

> The present theory, however, does not provide any simple set of empirically testable conditions which can be easily interpreted as a substantive theory. Furthermore, the general theory does not provide any constructive procedure for obtaining the desired numerical representation.

In fact, because of the difficulties, testable conditions have only been developed for simple models (Krantz, Luce, Suppes, and Tversky, 1972; Fishburn, 1973). Tversky (1967a) also leaves open the problem of "the development of appropriate error theories together with a statistical analysis of the problems of goodness-of-fit of the data to models."

Some of these problems will be addressed in the next chapter, relative to a specific model applied to the assessment of utilities in a multiple-attribute space.

Chapter V. POLYNOMIAL UTILITY ASSESSMENT

Section 1. Introduction

In this chapter, we outline how the theories and methods of multidimensional scaling and conjoint measurement may be employed to obtain one extension of the uniattribute utility assessment procedure of Novick and Lindley to multiple attributes. The process is straight-forward, and therefore it is not be discussed in great detail. Some of the interpretations of the resulting representation are interesting, however, and are reviewed.

In Section 2, the Novick and Lindley procedure is represented as a general conjoint measurement model, as formulated by F. W. Young (1972). It is then shown that this representation can be extended to multiple attributes in a straight-forward manner. The resulting representation is a general, fixed-state, discrete, multiattribute utility function.

This model requires that utilities be inferred for every point in the multiattribute space that is used in the elicitation procedure. In high-dimensional problems, this involves many parameters; therefore, the polynomial conjoint measurement theory is invoked to provide a more parsimonious representation. In particular, we look at a fourth-degree polynomial of the uniattribute scale representation. It should be remembered, however, that there is no good measure of the adequacy of the representation nor is there a theory of error of measurement for the general polynomial representation. These issues are discussed in Section 3.

By invoking techniques from multidimensional scaling, we can relieve some of the difficulties that led Novick and Lindley to infer utilities in the logodds metric. We suggest that the elicited indifference probabilities may be scaled before fitting the polynomial model of the utilities.

We also address the problem of the relationship between the model space (e.g., the resulting uniattribute scales in the polynomial representation) and the underlying object space. In particular, we suggest that a system of fourth-degree polynomial functions may be fit to the distortion between the two spaces when the underlying object space is assumed to be a subset of the product space of real numbers. The works of other authors along these same lines is cited.

In Section 3, we discuss the usefulness of the representation derived in this chapter and look at some of the difficulties. Several submodels, special cases of the general model, are presented and their interesting interpretations discussed. The problem of adequacy of the representation is also discussed.

Section 2. The Extended Novick-Lindley Procedure

Recalling the development of the Novick and Lindley fixed-state utility assessment procedure from Chapter 2, the basic model may be represented as

$$\text{(V.2.1)} \quad \log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \log\left(\frac{u_j-u_i}{u_k-u_j}\right)$$

where p_{ijk} is the probability that equates utility u_j with a gamble involving two states with utilities u_i and u_k, assuming that $u_k > u_j > u_i$. This will now be related to the general conjoint measurement model of F. W. Young (1972). First, however, it must be noted that we are concerned here with three variables (i, j, and k) whereas Young's formulation, using conventional matrix notation, was concerned with only two variables. This poses no real problem and be resolved any one of several ways. The easiest, perhaps, is to conceive of k indexing the columns of Young's matrices and i and j together indexing the rows.

We can therefore derive a convenient correspondence between the two models, recalling Young's general model:

$$\text{(V.2.2)} \quad h(x) = h_1(h_2(x_{i'j'}, \ldots, \text{[illegible]}))$$

by letting X be the matrix of inferred utilities (a column vector), i'=(i,j) and j'=k, h be an identity function, and

$$\text{(V.2.3)} \quad h_1(u_{i'}, u_{j'}) = h_1'(u_i, u_j, u_k) = \text{[illegible]}.$$

Now it must be noted that in the above formulation, u_i, u_j, and u_k are used to denote function values, in particular $u(x_i)$, $u(x_j)$, and $u(x_k)$, respectively, where x_i, x_j, and x_k are points, or objects, in the underlying uniattribute space. The generalization to multiple attributes is now clear. If we let $\underline{x}_i$, $\underline{x}_j$, and $\underline{x}_k$ be points in a multidimensional space and $u(\underline{x})$ be a function from the multidimensional space, of which $\underline{x}$ is a generic element, into the real numbers, then the above model (V.2.1) gives us a utility function over a multidimensional outcome space.

Some of the implications of this model should be pointed out. First, the gambles that are used in the elicitation procedure, which is exactly the same as advocated by Novick and Lindley and implemented by Novick, et alia (Isaacs and Novick, 1978; Novick, Hamer, Libby, Chen, and Woodworth, 1980), involves each point in the underlying space without regard to structure other than order in the utility metric. We make the same assumption that the multiple attribute utility function is a monotonically increasing function of the underlying dimensions. The resulting representation is a monotonic point (i.e., discrete) function with a vector argument. This is directly analogous to the implementation by Novick, et alia, cited above.

Because of the formulation of the model, every point in the underlying space must be involved in a gamble in order that a utility for every point be inferrable. There may be a large number of points in the space. Suppose, for example, that we wish to infer utilities in a three-dimensional space. We might then choose, for example, nine points of interest along each of the three dimensions. If we wish to view the complete product space of these three dimensions, we have 9*9*9 = 729 points in the space, and therefore 727 utilities to infer! (Recall that we assume that the least preferred point has utility zero and the most preferred point has utility one.)

As the number of utilities to be estimated is so large, the number of gambles for which indifference probabilities must be elicited is also large. We would like to present many more gambles than parameters, thus overfitting the model and allowing for coherence checking. We expect that this straight-forward extension of the Novick and Lindley procedure will be tedious for the decision maker and therefore not feasible in most cases.

Using the results of the theories of polynomial conjoint measurement, we do have an appealing alternative. If we are willing to assume a polynomial representation of the overall multiattribute utility function in terms of the marginal utility functions, we can achieve some parsimony. This gain in parsimony is not without cost, however; the cost is discussed in Section 3.

As an example, we might suppose that a fourth-degree polynomial representation would be sufficient to adequately approximate the overall, multiattribute utility function. The model would then consist of the weighted sum of all the powers and crossproducts of the marginal, uniattribute utility functions. The marginal utility functions, one along each dimension in the utility space, would be inferred from the data elicited.

The polynomial function must be constrained, since we are constraining our utility functions, the overall utility function and the marginal utility functions, to range from zero to one. With our assumption concerning the utilities of the least and the most preferred states, and assuming that every point in the product space is attainable (a relaxable assumption), the weights in the polynomial representation must be greater than or equal to zero, be less than or equal to one, and sum to one. These constraints are not particularly difficult and can easily be implemented with well-known numerical methods.

To see the parsimony of the model, we must determine the number of parameters that must be inferred. If we wish to represent the same three-dimensional model used as an example above, we have three uniattribute utility functions of nine states each and thirty-four weights in the polynomial function. This gives fifty-five parameters to infer (recall that two utilities in each of the three utility point functions are fixed). Thus the polynomial representation requires considerably fewer parameters (55 compared to 727)!

To implement the polynomial representation, we refer back to Young's general conjoint measurement model (equation V.2.2). In the formulation of the Novick and Lindley model, we observed that the function $h_2(.)$ was an identity function. By using a fourth-degree polynomial function for $h_2(.)$ instead, we get our desired representation.

Now we want to look at the metric in which we use the indifference probabilities to infer the utilities. This is a concern regardless of the representation we choose for the utilities. In the original Novick and Lindley formulation, the logodds metric was selected because it was thought that the subjective measurement in the tails of the domain (i.e., indifference probabilities near zero or near one) would be more precise than in the middle of the domain so an adjusting transformation was needed. Recent research has shown that this is not necessarily the case (Mayekawa, 1981).

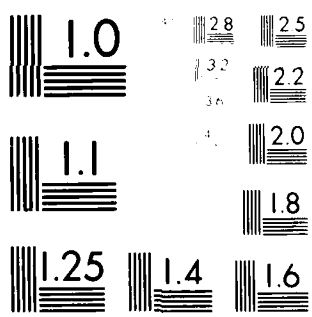
The use of any transformation (e.g., logodds, arcsine of the square root) for such purposes can be viewed as a scaling operation (see Luce and Tukey, 1964). For the reasons cited in the Luce and Tukey paper, i.e., that with the use of a particular transformation one does not know whether a different transformation might have given different results, the use of a general monotone transformation might be advisable.

Finally, we address the problem of relating the inferred utility space with the underlying object space. If the underlying object space is in fact discrete and utilities have been exhaustively inferred by the procedure advocated above, there is nothing more to do. It [illegible] the case, however, particularly in educational applications

AD-A098 230 IOWA UNIV IOWA CITY DIV OF EDUCATIONAL PSYCHOLOGY F/G 12/1
MULTIATTRIBUTE FIXED-STATE UTILITY ASSESSMENT.(U)
MAR 81 D L LIBBY, M R NOVICK N00014-77-C-0428
UNCLASSIFIED TR-81-4 NL

2 OF 3
AD A
098230

982



that the underlying object space is a subset of the product space of the real numbers (e.g., each dimension in the space is a subset of the real numbers). In such cases, it is useful to construct a functional relationship between the derived utility space and the underlying space.

In the unidimensional case, examples of this construction may be found in Novick, et alia (Isaacs and Novick, 1978; Novick and Lindley, 1979; Novick, Hamer, Libby, Chen, and Woodworth, 1980), where pieces of probability functions are fit to the utility point function. In the multidimensional case, examples may be found in Ward, et alia (Ward, 1977; Hendrix, Ward, Pina, and Haney, 1979; Ward, Pina, Fast, and Roberts, 1979), where techniques of "policy capturing" and "policy specifying" are used. Policy capturing involves predicting the behavior of a judge with relevant variables. This is also called "bootstrapping" (see Dawes, 1973). Policy specifying involves logically specifying functional relationships among relevant variables in a hierarchical, binary tree fashion to develop a model.

Presumably, one could fit pieces of probability functions to the marginal utility functions derived above. A simpler method would be to employ a technique similar to the bootstrapping in Ward, et alia. Instead of using the utility estimates specified by a judge, however, one might infer the marginal utilities from the behavior of a judge in the manner described above and then fit polynomial functions of the variables in the underlying space to the marginal utility functions.

Thus we have the following two-stage model of utilities. Let $\underline{\theta}$ be an arbitrary point in the underlying space. Then

(V.2.4) $u(\underline{\theta}) = p_4(\underline{u}(\underline{\theta}))$

where $p_4(\cdot)$ is a fourth-degree polynomial of its arguments, with crossproducts, and

(V.2.5) $\underline{u}(\underline{\theta}) = \underline{q}_4(\underline{\theta})$

where $\underline{q}_4(\cdot)$ is a system (vector) of fourth-degree polynomial functions of the components of $\underline{\theta}$. Note that each of the components of the vector $\underline{u}(\cdot)$, the marginal utility functions, is a polynomial function of all the components, and crossproducts, of the underlying space. This complexity gives us considerable flexibility in mapping the distortion from the underlying object space to the utility space.

Implications of this model and a couple of submodels are discussed in the next section.

Section 3. Discussion of the Polynomial Model

As we showed in Section 2, the polynomial representation can provide us with a more parsimonious model than the point function representation covering the entire underlying object space. Parsimony not only give us a conceptually simpler model, but also one that is easier to estimate. Because there are fewer parameters for which to infer values, our elicitation, or data collection, task is much simpler and therefore less tedious to both the the measurer (which may be a computer) and the decision maker (from whom the data are elicited).

The polynomial model is not without costs, however, as intimated above. The major costs are (a) it is an approximation except in strictly defined situations, (b) there are no tests for correctness except in a few special cases, and (c) there is no theory of error or distribution.

The strength of the model, as stated by Tversky (1967a) and quoted in Chapter IV, is that a polynomial function is very versatile. A polynomial function can be found to approximate any bounded, continuous surface. It is still an approximation and therefore does not fit exactly, in general. Thus, it becomes important to have testable conditions of correctness of the model, to identify those cases when the fit is exact, or a strong error theory to support and evaluate the approximation.

Except for simple cases, there are no testable conditions for the general polynomial model (Tversky, 1967a; Krantz, Luce, Suppes, and Tversky, 1972). Two of the special cases were discussed in Chapter IV, the additive and the multiplicative representations. Cancellation conditions and independence conditions for these and other simple models have been studied extensively in the literature (e.g., Fishburn, 1965; Keeney, 1971; Krantz, Luce, Suppes, and Tversky, 1972; Narens, 1974; Keeney and Raiffa, 1976). Conditions for the appropriateness of certain simple distributive models have been investigated (Krantz, Luce, Suppes, and Tversky, 1972). Conditions for simple bilinear forms have also been studied (Fishburn, 1973b). These models are all simple and at least suffer from the flaw that the conditional utilities, those derived from fixing all but one attribute at some combination of levels, are all linear functions of the marginal utility of the remaining variable attribute.

The existing theories of measurement giving testable conditions for correctness of the model also suffer from an assumption of infallible data. There exists no theory of error or of distribution. If one of the axioms, say the transitivity part of the weak ordering axiom, fails then the model fails. The failure of transitivity could be due to error of measurement, but there is no way of accounting for this in the current theories of conjoint measurement.

One approximation to a solution to the problem of model selection with fallible data is a procedure like that used in regression analysis or the analysis of variance. As indicated in a previous chapter, classical regression analysis and analysis of variance can be considered as conjoint measurement problems with strong metric assumptions. One could infer the overall multiattribute utility function and, using it as the dependent variable, apply one of the model selection algorithms of classical regression analysis to the proposed model.

There are two major difficulties with this approach, however. One is again concerned with the lack of a theory of error or distribution. The dependent variable not only has no established theory of error but is also inferred, not observed. Secondly, the procedure requires that a large number of parameters be inferred. Not only must the entire overall utility point function be inferred, but also all the parameters in the model. We do end with a more parsimonious model, but a considerable data collection task is required.

Finally, the proposed polynomial model is not always sufficiently more parsimonious than the general point function to be worth the effort. The example selected above to illustrate the polynomial model cast it in a good light. Indeed, when one is interested in a three-dimensional or higher representation with many points along each dimension, the polynomial representation will be more parsimonious. When one is interested in a two-dimensional representation, the advantage of the polynomial model disappears. In the three-dimensional case, the general polynomial model becomes attractive only if we are modelling more than five points along each dimension.

Since many of the applications of interest are likely to involve models of low dimensionality, generally two or three because of the conceptual simplicity, the polynomial representation may be marginally useful in practice.

The general polynomial formulation does have two submodels that are of some interpretational interest. First, if in equation (V.2.5) the system of polynomial equations $\underline{q}_4(\cdot)$ degenerates into a single polynomial function $q_4(\cdot)$, and equation (V.2.4) is taken to be an identity function, we have a polynomial fit of the variables of the underlying object space to the overall utility point function. This is essentially the "policy capturing" paradigm of Ward, et alia, cited above.

Second, if we let each equation in the system of polynomial equations (V.2.5) be a function of only one argument, i.e., involve only one of the dimensions in the underlying object space, we get a simpler model. Note, however, the implications of this model. In the general polynomial model, we allow any polynomial rotational distortion between the underlying space and the utility space to be modelled. It is even possible for the underlying space to be collapsed into a utility space of fewer dimensions (or expanded into more dimensions). This second submodel assumes a rotational restriction. It is assumed that the dimensions of the utility space are the same as those for the underlying object space and it is merely the scale on each axis that is distorted.

These two submodels are conceptually interesting in certain situations but they do not address the major problems in the application of the polynomial model. The problems of correctness of the model and goodness of fit still remain.

Chapter VI CONDITIONAL EXPECTED UTILITY ASSESSMENT

Section 1. Introduction

In this chapter we will present two axiomatizations of conditional subjective expected utility and use them to justify the method of multiattribute utility assessment that is the goal of this paper. The emphasis on conditional subjective expected utility theory is in response to criticisms of the axiomatizations of von Neumann and Morgenstern and of Savage, presented in Chapter II, which are both unconditional theories.

The unconditional theories define as primitive elements sets that represent consequences, acts, and states of the world. It is assumed that the states of the world and the relationships between them and the consequences are unaffected by the choice of the acts. In other words, it is assumed that one and only one state of the world is true and this fact is unalterable by the choice of an act by the decision maker.

There have been several criticism of this assumption in the literature (Luce and Krantz, 1971; Fishburn, 1973c;

Fishburn, 1974; Balch, 1974; Balch and Fishburn, 1974). Briefly, these objections center on the unconditional nature of the theory (i.e., that the choice of act for implementation can have no effect on the state of the world obtained) and the proposition of 'constant acts' (i.e., acts that produce the same consequence regardless of the true state of the world).

In response to these objections, two theories of conditional subjective expected utility have been developed (Luce and Krantz, 1971; Fishburn, 1973c). These two axiomatizations are presented here for reference, the Luce and Krantz axiomatization in Section 2 and the Fishburn axiomatization in Section 3.

In the Luce and Krantz theory, the primitive elements are events, or states of the world, consequences, and acts, which are functions from the set of events into the set of consequences. The acts are conditional, on subsets of the set of events, and the theory provides for a measurable utility and conditional subjective probabilities on these conditional acts. The authors prove an additional theorem that provides a measurable utility on the set of consequences. The application of the theory is restricted, however, to those sets that include all subsets and finite unions of disjoint subsets of the set of events of interest.

In the Fishburn theory, there is no such restriction on the set of events, but the set of acts is more restrictive, being a "mixture set" of all simple acts, those that obtain a single consequence (Krantz and Luce, 1974). It is this enrichment of the set of available acts that allows the theory to be applied to a less restrictive set of states of the world.

In Section 4, these theories are discussed in terms of their applicability to the situation under consideration in this paper, namely multiattribute utility assessment. Of particular interest is the manner in which each of the theories avoids involving consequences. Although each theory may be formulated in terms of consequences -- Luce and Krantz (1971) prove a secondary theorem and Fishburn's act-state pairs (1973c) may be called consequences without loss of generality (Pratt, 1974) -- both avoid bringing consequences into the theory at an axiomatic level where they might provide intuitive appeal. In this regard, comments by Jeffrey (1974) are cited that motivate a "new" axiomatization of conditional subjective expected utility.

In addition, these formulations allow us to address the problem of structure in a multiattribute situation. We show that uniattribute utility functions may be constructed that are conditional expected utilities of the overall utility function. We also show that the Novick and Lindley fixed-state utility assessment procedure may be viewed as a conditional subjective expected utility procedure. If we view the utility function as if it were a probability measure, some interesting interpretations develop.

Section 2. The Luce and Krantz Axiomatization

In this section, we present an axiomatization that leads to a simultaneous measurement of utility and subjective probability (Luce and Krantz, 1971). The axioms are similar in nature to those in the previously presented axiomatizations, and the resulting representation is similar though specifically formulated for the application to utility assessment.

The intention of the axiomatization is to provide a theoretical measurement foundation for utility assessment formulated in a more realistic way than the axiomatizations of utility developed by von Neumann and Morgenstern (1947) and by Savage (1954). The difficulty with these two previous axiomatizations, according to Luce and Krantz, is their representation of utility unconditionally. Specifically, both previous axiomatizations of utility require that the states of nature can no way be altered by the acts chosen by the decision maker. In most cases, Luce and Krantz argue, the decisions that we as decision makers consider limit the domain of possible events, thus altering our perceived subjective probability of the occurrences of those events as well as our perceived utilities for the outcomes.

The axioms proposed by Luce and Krantz are presented in Table VI.2.1. As stated by the above authors:

> They are mostly not transparent at first sight, but they become reasonably compelling as normative principles once their meanings are grasped.

Axioms 1, 8, and 9 are structural, guaranteeing that the sets under consideration are sufficiently rich to provide a basis for applying the other axioms and to insure a solution to certain equivalence relations. The other axioms, 2 through 7, are described as "rational preference behavior axioms" by Luce and Krantz. These are similar to axioms in other axiomatizations of utility that are normative (what one should do) rather than descriptive (what one does).

Axiom 1 insures that all possible combinations of conditional decisions are represented in the set of decisions under consideration. Note that some of these may be not meaningful. As Luce and Krantz state:

> To apply our theory to real-world decisions, we must therefore suppose that "natural" decisions, such as f_A and g_B (in Table VI.2.1), are enriched by certain artificial ones.... The measurer must

Table VI.2.1

The Luce-Krantz Axiomatization

(Luce and Krantz 1971)

Let $\mathcal{E}$ be an algebra of events (i. e., closed under unions and compliments) of subsets of a given set $\mathcal{X}$ of possible chance events; let $\mathcal{N}$ be null events, a subset of $\mathcal{E}$, including at least the empty set $\emptyset$, that is characterized by the axioms and that will be the events assigned probability zero in the representation; let $\mathcal{C}$ be consequences, an arbitrary set; let $\mathcal{D}$ be conditional decisions, a set of functions from non-null events ($\mathcal{E}$-$\mathcal{N}$) into $\mathcal{C}$; and let $\geq$ be a preference ordering, a binary relation over $\mathcal{D}$.

Definition 1: $< \mathcal{X}, \mathcal{E}, \mathcal{N}, \mathcal{C}, \mathcal{D}, \geq >$ is a conditional decision structure if for all A, B in $\mathcal{E}$-$\mathcal{N}$, R, S, in $\mathcal{E}$, and all f_A, f_A^i, f_{AUB}, f_{AUR}, g_B, g_B^i, h_A^i, k_B^i in $\mathcal{D}$, the following nine axioms are satisfied.

Axiom 1: (i) if $A \cap B = \emptyset$, then $f_A \cup g_B$ is in $\mathcal{D}$; (ii) if $B \subset A$, then f_A is restricted to B in $\mathcal{D}$.

Axiom 2: $\geq$ is a weak ordering of $\mathcal{D}$.

Axiom 3: If $A \cap B = \emptyset$ and $f_A \sim g_B$, then $f_A \cup g_B \sim f_A$.

<u>Table VI.2.1</u>

(continued)

<u>Axiom 4</u>: If $A \cap B = \emptyset$, then $f_A^1 \geq f_A^2$ if and only if $f_A^1 \cup g_B \geq f_A^2 \cup g_B$.

<u>Axiom 5</u>: If $A \cap B = \emptyset$, $f_A^i \sim g_B^i$, $i=1, 2, 3, 4$, $f_A^1 \cup k_B^1 \sim f_A^2 \cup k_B^2$, and $h_A^1 \cup g_B^1 \sim h_A^2 \cup g_B^2$, then $f_A^3 \cup k_B^1 \geq f_A^4 \cup k_B^2$ if and only if $h_A^1 \cup g_B^3 \geq h_A^2 \cup g_B^4$.

<u>Axiom 6</u>: If $A \cap B = \emptyset$, N is a sequence of consecutive integers, not $g_B^1 \sim g_B^2$, and $f_A^i \cup g_B^1 \sim f_A^{i+1} \cup g_B^2$ for i, $i+1$ in N, then either N is infinite or $\{f_A^i \mid i \text{ is in } N\}$ is unbounded.

<u>Axiom 7</u>: (i) If R is in N and $S \subset R$, then S is in N;
(ii) R is in N if and only if, for all f_{AUR} in D, $f_{AUR} \sim f_A$, where f_A is the restriction of f_{AUR} to A.

<u>Axiom 8</u>: (i) $E - N$ contains at least three pairwise disjoint elements; (ii) $D / \sim$ contains at least two distinct equivalence classes.

<u>Axiom 9</u>: (i) If A and g_B are given, then there exists h_A in D for which $h_A \sim g_B$; (ii) if $A \cap B = \emptyset$ and $h_A^1 \cup g_B \geq f_{AUB} \geq h_A^2 \cup g_B$, then there exists h_A in D such that $h_A \cup g_B \sim f_{AUB}$.

> be prepared to present for serious consideration by the decision maker some rather artificial alternatives, and the decision maker must be induced to make realistic decisions among them. The usual technique is to pose hypothetical questions.

Note that this is the practice of the Novick and Lindley procedure described in Chapter II.

Axiom 2 is analogous to axioms in the Luce-Tukey, the Roskies, and the Tversky systems. It means that the binary relation $\geq$ is reflexive and transitive. Axiom 3 simply means that if two decisions are equivalent in preference, then a decision that randomly provides one or the other is equivalent to either.

Axiom 4 states that if one decision is at least as preferable as another, then the preference ordering should remain the same when both are combined with some other, unrelated decision. Axiom 5 is the one that helps to guarantee that the resulting utility representation will attain at least an interval scale of measurement by postulating that there be no reversals in preference. Axiom 6 is an Archimedean axiom like those in other axiomatizations.

Axiom 7 is concerned with the events that will have zero probability in the resulting representation, guaranteeing that subsets of null events are themselves null and that null events have no influence on non-null decisions. Axiom 8 is another of the structural axioms, insuring that the sets in question have sufficiently many elements so that the axioms are applicable. The notation, $\mathcal{D}/\sim$, denotes the set of all equivalence classes derived from $\mathcal{D}$. Axiom 9 is a solvability axiom similar to those in other axiomatizations.

The primary result of the axiomatization is the following theorem:

<u>Theorem 1</u>: Suppose that $< \mathcal{X}, \mathcal{E}, \mathcal{N}, \mathcal{C}, \mathcal{D}, \geq >$ is a conditional structure in the sense of Definition 1. Then there exist real-valued functions u on $\mathcal{D}$ and P on $\mathcal{E}$ such that $< \mathcal{X}, \mathcal{E}, P >$ is a finitely additive probability space and, for all A, B in $\mathcal{E}-\mathcal{N}$, R in $\mathcal{E}$, and f_A, g_B in $\mathcal{D}$,

(i) R is in $\mathcal{N}$ if and only if $P(R) = 0$;
(ii) $f_A \geq g_B$ if and only if $u(f_A) \geq u(g_B)$;
(iii) if $A \cap B = \emptyset$, then
$$u(f_A \cup g_B) = u(f_A) P(A|A \cup B) + u(g_B) P(B|A \cup B).$$

Moreover, P is unique and u is unique up to a positive linear transformation.

It should be noted that this theorem does not assign a utility function to the set of consequences, but to the set of conditional decisions. Thus, as stated by Luce and Krantz, "they do not seem to fulfill our original goal and certainly they are different in this respect from all other theorems of expected utility." In particular, these refer to the axiomatizations by von Neumann and Morgenstern and by Savage. Another theorem is proved by Luce and Krantz that establishes a utility function defined on the consequences. Luce and Krantz are also able to show that, for the finite case, any conditional model with their representation can be restated in an equivalent unconditional form, and vice versa.

Section 3. The Fishburn Axiomatization

In this section we present a second axiomatization of conditional subjective expected utility (Fishburn, 1973c; Fishburn, 1974). The set of states of the world is generalized from that of Luce and Krantz, in that there is no requirement that all unions and subsets be included. To achieve the desired representation, however, the set of acts, or decisions or gambles, is enriched to be a "mixture set."

A mixture set is defined as follows (Herstein and Milnor, 1953):

> A set S is said to be a mixture set if for any a, b in S and for any p we can associate another element, which we write as pa + (1-p)b, which is again in S, and where
>
> $$1a + (1-1)a = a,$$
> $$pa + (1-p)b = (1-p)b + pa,$$
> $$q(pa + (1-p)b) + (1-q)b = (qp)a + (1-qp)b,$$
>
> for all a, b in S and all q, p.

The resulting representation is not substantially different from that of Luce and Krantz (1974). One representation does not dominate the other so that the choice must be dependent on the situation or sets under consideration or those likely to occur.

Table VI.3.1 presents the axioms of the Fishburn system. They are fully discussed in the original presentation (Fishburn, 1973c) and so will not be discussed in detail here. The Fishburn paper also cites situations and consequences of the failure of certain of the axioms. The implications of each of the axioms are briefly presented below.

Axiom (A1) means that the preference relation $>$ is asymmetric ($a>b \rightarrow$ not $b>a$) and negatively transitive (not $a>b$ & not $b>c \rightarrow$ not $a>c$). Defining "indifference" ($\sim$) as

$$a \sim b \text{ if and only if not } a>b \text{ \& not } b>a$$

and "preference-indifference" ($\geq$) as

$$a \geq b \text{ if and only if } a>b \text{ or } a \sim b,$$

it follows that $\sim$ is an equivalence relation (reflexive, symmetric, and transitive) and $\geq$ is transitive and complete.

Axiom (A2) is a continuity axiom, similar in intent to the Archimedean axioms of the systems presented in previous

Table VI.3.1

The Fishburn Axiomatization

(Fishburn, 1973c)

Definition 1. $(\mathcal{X}, \mathcal{E}, >)$ is an ordered algebraic mixture system if and only if $\mathcal{X}$ is a mixture set, $\mathcal{E}$ is a Boolean algebra of the set S of states of the world, $>$ is a binary relation on $\mathcal{X} \times \mathcal{E}'$ (where $\mathcal{E}' = \mathcal{E} - \{\emptyset\}$) and, for all A, B, C in $\mathcal{E}'$ and x, y, z, w in $\mathcal{X}$:

(A1) $>$ is a weak order,

(A2) $\{p: (px+(1-p)y)A \geq zB\}$ and $\{p: zB \geq (px+(1-p)y)A\}$ are closed (in the relative usual topology for [0,1]),

(A3) $xA \sim zB$ & $yA \sim wB \rightarrow (.5x+.5y)A \sim (.5z+.5w)B$,

(A4) $A \cap B = \emptyset$ & $xA \geq xB \rightarrow xA \geq x(A \cup B) \geq xB$,

(A5) $x > y$ for some x, y in $\mathcal{X}$,

(A6) $A \cap B = \emptyset \rightarrow xA > xB$ & $yB > yA$ for some x, y in $\mathcal{X}$,

(A7) If A, B, C are mutually disjoint and if there is an x in $\mathcal{X}$ such that $xA \sim xB$, then there is a y in $\mathcal{X}$ such that exactly two of yA, yB, and yC are indifferent.

where xA, for x in $\mathcal{X}$ and A in $\mathcal{E}'$, is written for the ordered pair (x,A) in $\mathcal{X} \times \mathcal{E}'$ and xS, for x in $\mathcal{X}$, is abbreviated x.

Table VI.3.1

(continued)

Theorem 1: Suppose that $(\mathcal{X}, \mathcal{E}, >)$ is an ordered algebraic mixture system. Then there is a real-valued function u on $\mathcal{X} x \mathcal{E}'$ and a finitely-additive probability measure P_A on $\{A \cap B$: B in $\mathcal{E}\}$ for each A in $\mathcal{E}'$ such that (i) $xA > yB$ if and only if $u(xA) > u(yB)$ for all xA, yB in $\mathcal{X} x \mathcal{E}'$; (ii) $u(.A)$ is linear (as a function on $\mathcal{X}$) for each A in $\mathcal{E}'$; (iii) $P_C(A) = P_C(B)P_B(A)$ whenever $A \subseteq B \subseteq C$, A in $\mathcal{E}$, and B, C in $\mathcal{E}'$; (iv) $u(x(A \cup B)) = P_{A \cup B}(A)u(xA) + P_{A \cup B}(B)u(xB)$ whenever x in $\mathcal{X}$, A, B in $\mathcal{E}'$, and $A \cap B = \emptyset$. Moreover, a real-valued function v on $\mathcal{X} x \mathcal{E}'$ and finitely-additive probability measures Q_A on $\{A \cap B$: B in $\mathcal{E}\}$ for each A in $\mathcal{E}'$ satisfy (i) through (iv) in place of u and the P_A when these satisfy (i) through (iv) if and only if v is a positive linear transformation of u (if and only if there are numbers $a > 0$ and b such that $v(xA) = au(xA) + b$ for all xA in $\mathcal{X} x \mathcal{E}'$) and $Q_A \equiv P_A$ for each A in $\mathcal{E}'$.

chapters. One implication of this axiom is that the set is empty only if zB is not preferred (is preferred) to both xA and yA.

Axiom (A3) is crucial to provide the structure needed to derive the desired representation. It is a weak version of the sure-thing axiom found in other systems. Axiom (A4) is an averaging axiom that acts in the alignment of the utility functions in the resulting representation so that the utilities of all the elements will have the same, comparable scale. These two axioms may be difficult to accept on face value. Axiom (A3) is not very intuitively appealing, and Axiom (A4), though intuitively appealing, may be unacceptable because of the possible value of certainty.

Axiom (A5) simply states that there are at least two distinct elements, in terms of preference, in the set \, and Axiom (A6) states that the set \ is rich enough so that state A is preferred to state B for one act but the reverse is true for at least one other act. The implication of this latter axiom is that no state dominates any other state in preference.

Axiom (A7) is required to guarantee that the probabilities in the representation are additive. Fishburn gives an example where (A1) through (A6) are satisfied but (A7) fails and the probabilities are not additive.

Section 4. Conditional Expected Utility Assessment

The difference between the Fishburn axiomatization and the Luce-Krantz axiomatization was described as follows (Krantz and Luce, 1974, in comment to Balch and Fishburn, 1974):

> It seems to us that the chief difference between the two systems lies right at this point: the LK system imposes great richness on its outcome structure, but can get on with as few as three atomic events (or even two, with a little extra effort); the BF system can deal with any set of basic action alternatives, but utilizes the elaborate mixture-space apparatus. Presumably the latter apparatus could be made more qualitative by moving in the direction of the Savage axioms; but in any case, what is required is a very fine-grained structure of events or probabilities. The essence of this difference is familiar from the contrast between the utility measurement procedure of (Davidson, Suppes, and Siegel (1957) and Mosteller and Nogee (1951)). Does one best measure utility by trading off value and probability or by trading off value against value? The latter has more face validity and is more easily generalized to situations where the subjective probabilities are not well behaved; it is the method of Davidson et al. and of the LK system. The former method gets along with a much simpler structure of basic options; it is the method of von Neumann and Morgenstern, Mosteller and Nogee, Savage, and others, and is the basis of the BF system.

Both systems are usable in justifying the utility assessment procedure advocated by Novick and Lindley, and adopted for use in this paper. The Fishburn axiomatization is perhaps a more direct justification because of its common ties with the works of von Neumann and Morgenstern and of Savage.

The Luce and Krantz axiomatization provides a theoretical measurement foundation for the utility assessment procedure advocated by Novick and Lindley as follows. The Novick-Lindley procedure selects three states, assuming that they are ordered, and equates the expected utility of a gamble, involving the most preferred of the three states and the least preferred, with the utility of the third, middle preferred state. Although the equation (actually a system of such equations) can be solved unconditionally, it is convenient to view each equation conditionally. If we restrict our attention to the three states, and recalling that the utility function is unique only up to a positive linear transformation, we may assume *conditionally* that the least preferred state of the three has a utility of zero and the most preferred state has a utility of one. Thus, the conditional utility of the the third state, of middle preference, is equal to the (conditional) probability of the gamble.

In the Luce-Krantz axiomatization, a single state can be represented as a constant decision and a gamble, also a decision, as a finite union of constant decisions. Thus we can formulate the Novick-Lindley equations in terms of a gamble and a constant decision as follows: let A = {least preferred state}, B = {state of middle preference}, and C = {most preferred state}. Theorem 1 of the Luce-Krantz system allows us to conceive of the utility of the gamble involving the most and the least preferred states, a finite random union of sets A and C, as the sum of the utilities of each set multiplied by the subjective conditional probability of its occurrence. Then Axioms 9 and 3 allow us to equate this utility of the gamble to the utility of the third set B, representing the state of middle preference.

The Fishburn axiomatization also provides a theoretical foundation, as follows. As above, let A={state of least preference}, B={state of middle preference}, and C={state of most preference}. Assuming that the axioms hold, Theorem 1 concludes that

(VI.4.1) $u(w(A \cup C)) = P_{A \cup C}(A)u(wA) + P_{A \cup C}(C)u(wC)$

for some w in X. Axiom (A2) guarantees that

(VI.4.2) $w(A \cup C) = (px + (1-p)y)(A \cup C) \sim zB$

for some x, y, z in X. Thus, by Theorem 1 again, $u(w(A \cup C)) = u(zB)$.

Since both axiom systems can adequately justify the procedure adopted for this paper, the choice between them is one of convenience. Given that the application of interest (in educational selection and assignment) will likely satisfy the axioms of both systems, we could assume both, thus providing a very strong theoretical base for the elicitation procedure.

One criticism of both axiom systems is the absence of the set of consequences in the axioms and the main theorems. This is a minor fault, however, as both systems may be extended or reformulated to involve the set of consequences. Luce and Krantz (1971) state and prove a theorem that provides utilities on the set of consequences, as indicated in Section 2. In Fishburn's axiomatization, the act-state pairs can be conceived as consequences, as pointed out by Pratt (1974). Jeffrey (1974) suggests a "holistic" approach that includes consequences in the set of primitive elements. Indeed, states of the world, acts, and consequences all may be viewed as events.

Now that we have a firm theoretical justification for the elicitation procedure of Novick and Lindley, we turn our attention to the multiattribute situation. As in either axiom system, let S be the set of states of the world, and let S be multiattribute, i.e., $S = S_1 x S_2 x \ldots x S_k$. Similarly, let $\ell = \ell_1 x \ell_2 x \ldots x \ell_k$ be the Boolean algebra of the set S. For ease of demonstration, we will make some stronger assumptions than necessary, namely that each ℓ_i is an interval on the real line. The extension illustrated here is valid for other types of sets ℓ_i, such as subsets of the set of integers, because of the abstract nature of the sets defined in the axioms.

With these assumptions, the theorems of either system allow us to conclude, with suitable continuity extensions, that

$$\text{(VI.4.3)} \quad u(\underline{\theta}') = \int u(\underline{\theta}'')\,dP(\underline{\theta}''|\underline{\theta}') = E[u(\underline{\theta})|\underline{\theta}']$$

where $\underline{\theta}$ is in Θ, $\underline{\theta}'$ is in Θ', $\underline{\theta}''$ is in Θ'', and $\Theta' \times \Theta'' = \Theta$. Of particular interest is when $E' = E_i$ for some i. Thus we have identified a unidimensional utility function, $u(\underline{\theta})$, which Novick (1980) has called the **marginal utility function** of $\underline{\theta}'$, and have related it to the multiattribute utility function as its conditional expectation (given $\underline{\theta}'$). This is illustrated in Figure VI.4.1.

In addition to the marginal utility function, we define the **conditional utility function** as

$$\text{(VI.4.4)} \quad u(\underline{\theta}|\underline{\theta}^*) = \frac{u(\underline{\theta}) - \inf_{\underline{\theta}}(u(\underline{\theta})|\underline{\theta}^*)}{\sup_{\underline{\theta}}(u(\underline{\theta})|\underline{\theta}^*) - \inf_{\underline{\theta}}(u(\underline{\theta})|\underline{\theta}^*)}$$

where $\underline{\theta}$, $\underline{\theta}^*$ are in Θ. This is simply a rescaling of the multiattribute utility function $u(\underline{\theta})$ conditioned on $\underline{\theta}^*$ so that it ranges from zero to one. This is illustrated in Figure VI.4.2.

Figure VI.4.1

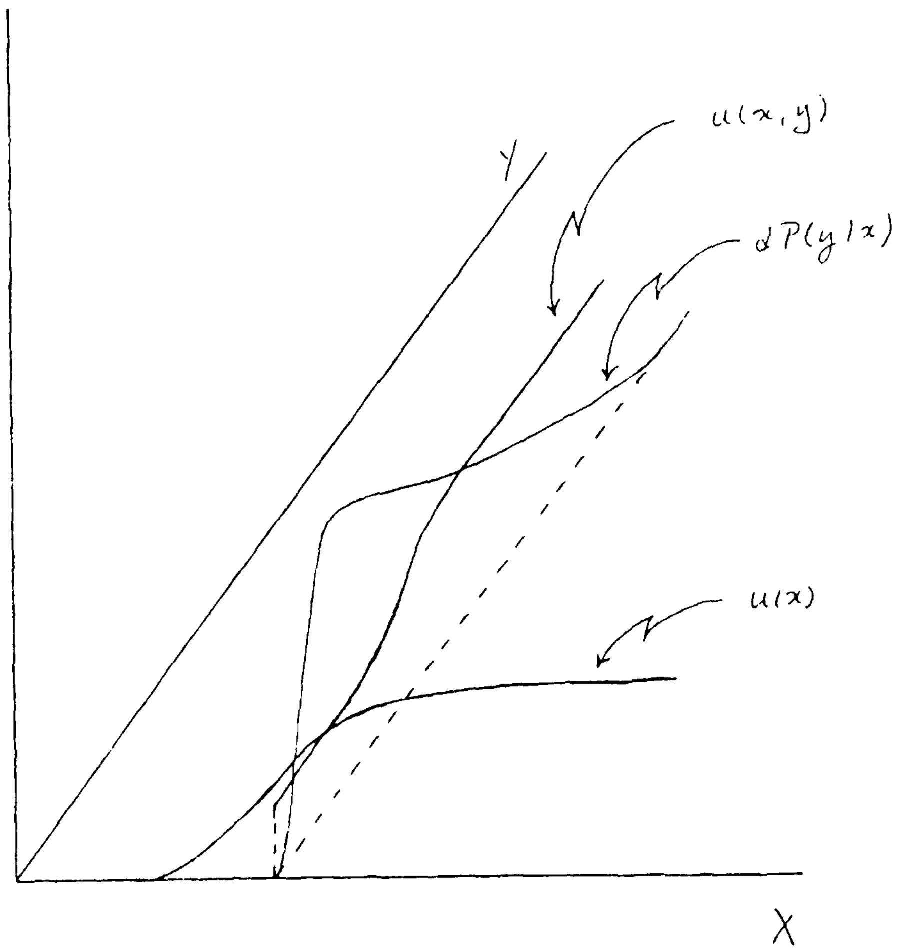
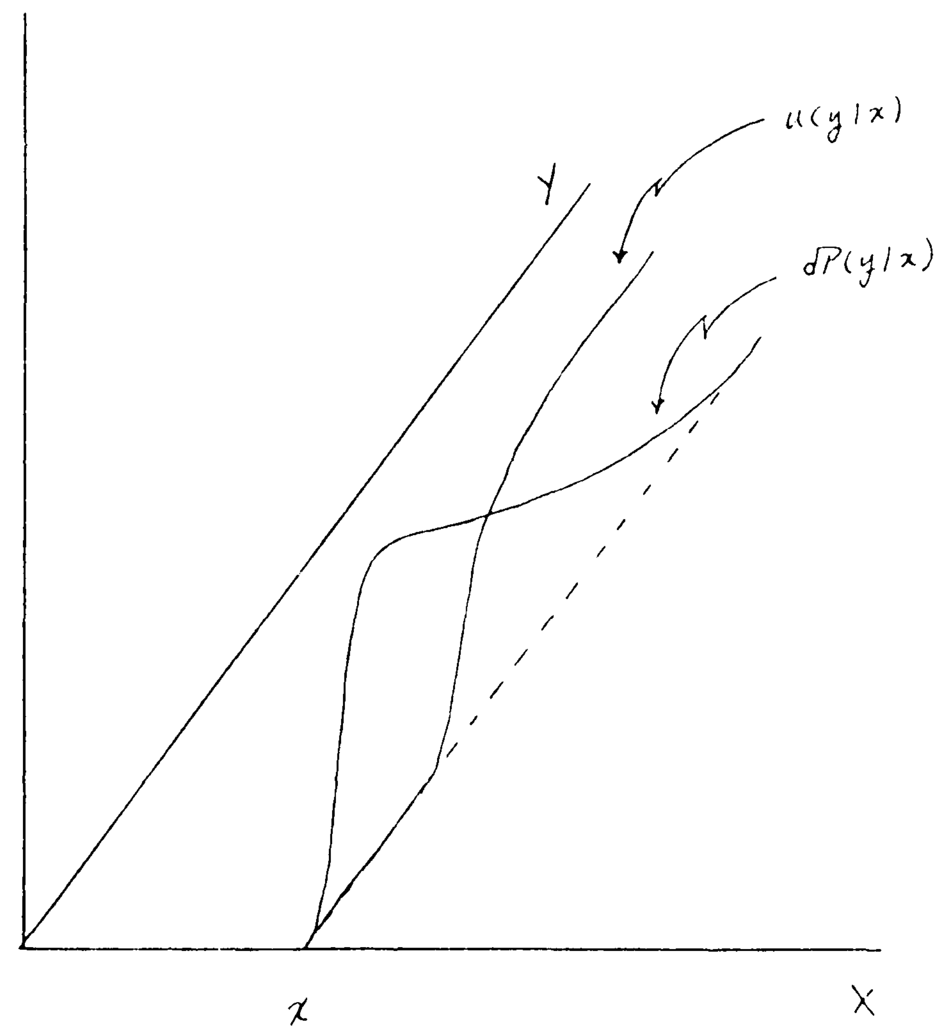


Figure VI.4.2



The conditional utility function is used in the Novick and Lindley theory as follows. Let $\underline{\theta}_i$, $\underline{\theta}_j$, $\underline{\theta}_k$ be three states ordered on preference such that $u(\underline{\theta}_i)<u(\underline{\theta}_j)<u(\underline{\theta}_k)$. The theory states that

$$\text{(VI.4.5)} \quad u(\underline{\theta}_j) = p\, u(\underline{\theta}_k) + (1-p)\, u(\underline{\theta}_i)$$

for some p, $0 < p < 1$. Then

$$\text{(VI.4.6)} \quad p = \frac{u(\underline{\theta}_j) - u(\underline{\theta}_i)}{u(\underline{\theta}_k) - u(\underline{\theta}_i)}$$

Thus, p is the conditional utility of $\underline{\theta}_j$, given that $\underline{\theta}_i$ is the least preferred state and $\underline{\theta}_k$ is the most preferred state under consideration, i.e., that $u(\underline{\theta}_i)=0$ and $u(\underline{\theta}_k)=1$.

The conditional utility function may be related to the marginal utility function by observing that the conditional expectation of the conditional utility function, $E(u(\underline{\theta}|\underline{\theta}^*)|\underline{\theta}')$, is equal to the rescaled marginal utility function:

$$\text{(VI.4.7)} \quad E\,[u(\underline{\theta}|\underline{\theta}^*)|\,\underline{\theta}'] = \frac{u(\underline{\theta}') \quad - \inf_{\underline{\theta}}(u(\underline{\theta})\,|\underline{\theta}^*)}{\sup_{\underline{\theta}}(u(\underline{\theta})\,|\underline{\theta}^*) - \inf_{\underline{\theta}}(u(\underline{\theta})\,|\underline{\theta}^*)}$$

noting that the restrictions $\underline{\theta}'$ are assumed to be a subset of the restrictions $\underline{\theta}^*$.

The three equations (VI.4.3), (VI.4.4), and (VI.4.7) provide a framework for the assessment of the multiattribute utility function $u(\underset{\sim}{\theta})$. As the marginal and the conditional utility functions are uniattribute utility functions, we may apply the elicitation procedure of Novick and Lindley. We may present choices between a gamble and a sure thing in any suitable manner. The resulting system of nonlinear equations in expected utility may be solved using any suitable numerical method. The multiattribute utility function may then be recovered by applying the relationships among it and the marginal and the conditional utility functions displayed above.

As developed in Novick and Lindley (1978), the use of convenient models for the utility function and the probability function over the states of interest can greatly simplify the computation of expected utilities. Those authors suggest using a cumulative distribution function for the utilities and conjugate probability functions. This suggestion works well in this situation. If we limit our attention to a two-attribute situation for the purpose of illustration, we can see that

(VI.4.8) $$u(\theta_1) = P\left[\tilde{\eta} < \tilde{\theta}_1, \tilde{\tau} < \theta_2 \mid \tilde{\theta}_1 = \theta_1\right]$$

and that

(VI.4.9) $$u(\theta_2 \mid \theta_1, \theta_2' < \theta_2 < \theta_2'') = P\left[\tilde{\eta} < \tilde{\theta}_1, \tilde{\tau} < \tilde{\theta}_2 \mid \tilde{\theta}_1 = \theta_1, \theta_2' < \tilde{\theta}_2 < \theta_2''\right]$$

With conveniently chosen utility and probability functions, the parameters of the resulting probability functions (for the marginal and conditional utility functions) may be calculated (Chen and Novick, 1981). This idea is developed more fully in the next chapter.

Chapter VII MULTIATTRIBUTE UTILITY ASSESSMENT

Section 1. Introduction

In this chapter, we outline a procedure to assess multiattribute utilities. Details of the procedure are generally referenced to other chapters of this paper or to the appendices. The sections of this chapter are divided as follows.

In this section, the paradigm for the multiattribute utility assessment is defined and the parameters of the problem are identified. The discussion is brief as the problem has been amply discussed in Chapters I and II of this paper. The design of the computer modules of the implementaton of the procedure developed in this paper is discussed in this and subsequent sections. The utility assessment procedure is presented in three phases: elicitation, coherence, and modelling.

In Section 2, the elicitation phase is discussed. Again, the discussion is brief because the procedures involved have been presented in detail elsewhere; for example, in Chapter II of this paper. Two elicitation procedures are implemented.

In the Conditional Coherence Procedure, a set of gambles for a conditional set of attributes is presented; the decision maker may then review and alter the indifference probability for any gamble in the conditional set before another conditional set is presented. Note that this procedure is analogous to the Global Coherence Procedure, otherwise known as the Least-Squares Coherence Procedure (Novick, Hamer, Libby, Chen, and Woodworth, 1980).

In the Regional Coherence Procedure, the decision maker is presented with additional coherence gambles after each pair of gambles from a conditional set. The procedure is like the procedure of the same name currently in the CADA Monitor (Novick, Hamer, Libby, Chen, and Woodworth, 1980).

In addition, two entry formats are available (Novick, 1980; Novick, Turner, and Novick, 1981). In one, the decision maker is presented with a structured, dynamic sequence of probabilities for each gamble triplet and responds whether the gamble is preferred, the sure thing is preferred, or the choice is indifferent. In the other format, the decision maker is presented with the same choice between a gamble and a sure thing and responds with the probability that would make the choice indifferent.

In Section 3, the coherence phase is discussed. The gamble triplets are presented along with the elicited indifference probabilities. If some estimation has taken place, the estimated, coherent indifference probabilities are presented, as well, along with the estimates of the parameters of the model chosen for the utilities. The decision maker may modify or delete any of the gamble triplets from the modelling phase. The decision maker may also request that additional gamble triplets be presented for coherence checking. This phase is entered after the modelling phase for acceptance of the estimates by the decision maker.

In Section 4, the modelling phase is discussed. A non-linear least-squares algorithm, based on the Newton-Raphson method, is used to estimate the parameters of the probability distribution chosen to model the utilities of the problem. Three utility models are implemented: the multivariate normal cumulative distribution function; the Dirichlet cumulative distribution function; and the multivariate generalized beta cumulative distribution function. The properties and the derivation of this latter distribution are presented in Appendix A; the former two distributions are well-known. The necessary derivatives of the three distributions are presented in Appendix B.

The scaling transformations of the coherence phase are proposed to be static and are therefore incorporated directly into the modelling phase. The modelling may be performed in one of three metrics: the metric of the indifference probabilities; the logodds metric; and the arcsine metric. The uses of the logodds and the arcsine metrics are well-known in the analysis and estimation of probabilities and proportions (e.g., Novick and Jackson, 1974). The necessary derivatives for these transformations are also presented in Appendix B.

Once the estimates have been calculated, the procedure returns to the coherence phase, where the decision maker may accept the estimates or modify the data for further estimation. A general discussion of the entire procedure is presented in Section 5.

In general, the problem may be defined as follows. One has a multiattribute decision to make and desires to assess the utility over this multidimensional space. We will let _ denote the attribute vector. We will assume that the attribute space is continuous, or at least that it may be adequately approximated by a continuous space. (This assumption is imposed by the choice of models. If a suitable multivariate discrete probability function could be postulated, it could be fit into this framework with only a little modification in the model estimation phase.)

The elicitation algorithm requires that we present gambles involving specific points within the attribute space; therefore, the decision maker must first be asked to determine these points. Several restrictions are useful.

First, the procedure is simplified if subsets of the points may be chosen so that the subsets are convenient conditional sets. The subsets are convenient if the conditional models for the subsets are simple models. Since we are modeling the utility functions by cumulative distribution functions, this convenience criterion requires that the conditional models be tractable cumulative distribution functions. Considering the models that we have selected, the most convenient set of points are those of a rectangular hypergrid. This allows the subsets to be states that differ in only one dimension, and the conditional models to be simple cumulative distribution functions.

This restriction also allows the simplest method for selecting the points to be used. The decision maker merely determines the number of dimensions in the problem and then selects several points along each dimension. These points determine the rectangular hypergrid.

This is not a substantial restriction. Although it is reasonable that some of the resulting multidimensional points may be improbable or difficult for the decision maker to value, the model for the utility function does not require that every point in the hypergrid be involved in the elicitation phase. This is an advantage of the functional model as opposed to the point estimation model, such as the least-squares procedure of the CADA Monitor (Novick, Hamer, Libby, Chen, and Woodworth, 1980). If a particular point poses a difficulty, the decision maker may simply choose to not judge any gamble triple that involves it.

Second, we require that the utility function be monotonically increasing in all dimensions. Again, this is a convenience for our choice of models; indeed, the models impose this condition. It is, however, a reasonable restriction. In many cases where this restriction is not met, the problem may be redefined.

For example, in a one-dimensional problem in which the utility is high in the middle of the attribute space and falls off towards both extremes, such as in some medical situations in which both a high dosage and a low dosage have low value, the attribute space may be "folded" about the modal point to conform.

Third, we assume that the utility function is bounded (Novick and Lindley, 1979), both above and below, and that the decision maker can select points that are of essentially minimum and maximum value. Thus we can assume, without loss of generality that the utility function ranges from 0 to 1.

This is in effect defining a conditional utility function. The point where all dimensions have the least value is assumed to have value 0, and the point where all dimensions have the greatest value is assumed to have value 1. With some additional information or assumptions, we could unconditionalize the resulting utility function; we will not, however, discuss this point further.

Section 2. The Elicitation Phase

The elicitation phase for the multiattribute utility assessment is essentially the same as was discussed in Chapter II and implemented in the CADA-1980 Monitor (Novick, Hamer, Libby, Chen, and Woodworth 1980). There are two elicitation algorithms: the Conditional Coherence Procedure and the Regional Coherence Procedure (Novick, Chuang, and DeKeyrel, 1979; Novick, Hamer, Libby, Chen, and Woodworth, 1980). There are two entry formats: direct entry (of the indifference probability) and "ends-in" (Novick, 1980; Novick, Turner, and Novick, 1981). In addition, the decision maker has the option of passing, or not making, a judgment.

In both the elicitation algorithms, the decision maker is presented with a sequence of gamble triples. Each triple consists of a gamble between two points in the attribute space and a "sure thing" point. The triples are chosen so that the sure thing point is intermediate in utility relative to the two points in the gamble and so that only one dimension in the attribute space is varying among the three points. The goal is the probability of the higher-valued point in the gamble such that the decision maker is indifferent between the gamble and the sure thing.

For example, suppose that we have three points in the attribute space, θ_i, θ_j, and θ_k in order of preference. The decision maker would be presented with a choice (hypothetically) between obtaining state θ_j for sure and gambling on obtaining state θ_k with probability p or obtaining state θ_i with probability 1-p. Our goal is to elicit the probability p such that the decision maker is indifferent in this choice.

In the Conditional Coherence Procedure, the gamble triples are presented in sets, where only one dimension of the attribute space is varying in each set. The procedure is like the least-squares procedure of the CADA-1980 Monitor, except that we are considering the one-dimensional, conditional utility function within each set. After judging each set, the decision maker may elect to review the judgments of that set and correct any errors.

The conditional sets, as they assume fixed states for all but one attribute, are chosen so that they cover the attribute space. The decision maker may elect to pass on any set. Within each set, the gamble triples are chosen so that the three states are adjacent in the dimension that is varying.

For example, suppose that the decision maker has chosen N states, θ_1 , ..., θ_N, along a single attribute dimension for elicitation. The decision maker would be presented with a choice between a state θ_i for sure and a gamble involving states θ_{i+1} and θ_{i-1} , where i is some integer between 1 and N.

The Regional Coherence Procedure differs only slightly from the above. As in the Conditional Coherence Procedure, the gamble triples are presented in conditional sets, where the states of all but one attribute are fixed. Again, the adjacent gamble triples are used.

The triples are presented in pairs, with each pair involving only four states. After each pair of adjacent triples is presented, the two other triples from the four states are presented with the implied indifference probabilities. For example, let θ_a, θ_b, θ_c, and θ_d be four adjacent states in a uniattribute assessment. The decision maker would be presented first with the two choices:

	for sure	gamble p	gamble $1-p$
1.	θ_b	θ_c	θ_a
2.	θ_c	θ_d	θ_b

and might respond with probabilities p_1 and p_2, respectively, that make the choice indifferent.

Based on this information, coherent indifference probabilities may be calculated for two other choices:

	for sure	gamble p	gamble $1-p$
3.	θ_b	θ_d	θ_a
4.	θ_c	θ_d	θ_a

This is accomplished by examining the conditional utility function over those four states. Since two utilities are fixed (at 0 and 1), only two gamble triples are needed to determine the conditional utility function. The decision maker may then accept the implied indifference probabilities or correct any of them. When they are acceptable, another pair of adjacent triples from the same conditional set is presented.

The above procedure is like the procedure of the same name in the CADA-1980 Monitor. When all the pairs from one conditional set have been presented, the decision maker may review the judgments and proceed on to the next set.

The entry formats are available in both of the above procedures. The direct entry format allows the decision maker to specify the indifference probability for the presented gamble triple. The "ends-in" format allows the decision maker to respond with a preference for the sure thing, a preference for the gamble, or indifference. When the decision maker is indifferent, another gamble triple is presented.

For example, suppose that the decision maker has chosen states θ_{i-1}, θ_i, and θ_{i+1} , in order of preference, for elicitation in a uniattribute decision problem. In the direct entry format, the decision maker would be presented with the following choice:

For Sure	Gamble p	1-p	P that makes you indifferent
θ_i	θ_{i+1}	θ_{i-1}	?

The decision maker would enter the probability for the gamble that would make the choice between the gamble and the for-sure state indifferent.

In the "ends-in" format, the decision maker is presented with a probability for the gamble as well as the gamble triple. The probability is initially either 0.1 or

0.9. For example, the decision maker might be presented with the following choice:

```
------------------------------------  Option:
I                                  I
I                  p  θ            I  0. Indifferent
I For sure θ           i+1  Gamble I  1. For sure
I           i     1-p θ            I  2. Gamble
I                      i-1         I  3. Restart
------------------------------------
```

Which would you prefer if the p = .10?

The response of the decision maker determines whether the indifference probability is higher than, lower than, or equal to the presented probability. The next probability to be presented is calculated to be near the opposite end of the resulting interval.

For example, suppose that the presented probability is 0.1 and the decision maker prefers the sure thing. Thus, we know that the indifference probability is in the interval (0.1,1.0). The next probability is calculated to be near the end of the interval opposite to the presented probability, or 0.9. If the decision maker then prefers the gamble, then we know that the indifference probability is in the interval (0.1,0.9). This sequence continues until the decision maker responds with indifference or until the interval becomes insignificantly small; this threshold is set to 0.05. In this latter case, the middle of the interval is selected.

In both options, the decision maker may avoid judging the gamble triple. In the "ends-in" option, the decision maker may also elect to begin the questioning for the current gamble triple again.

Section 3. The Coherence Phase

In actuality, the previous section described the iteration between the elicitation phase and the coherence phase. In the Conditional Coherence Procedure, coherence may be checked after each conditional set of gamble triples is presented. In general, however, there will be insufficient data to make initial utility estimates; thus, the coherence checking involves only the review of the assessed indifference probabilities.

In the Regional Coherence Procedure, coherence is checked after each pair of gamble triples is presented. There is generally insufficient data to make utility estimates for the entire conditional set, but it is possible to calculate the indifference probabilities for gambles in the region.

The rest of the multiattribute utility assessment procedure is primarily an iteration between the coherence phase and the modelling phase. After the parameters of the model for the utility function have been estimated, the implied indifference probabilities are presented along with the assessed indifference probabilities. The decision maker may decide to accept the fit, modify any of the assessed indifference probabilities, or delete any of the gamble triples from consideration.

The indifference probabilities, both assessed and implied, are presented by conditional set. The difference between the two, in the selected metric, is shown, as are the states involved in the gamble triple.

Optionally, the decision maker may view the estimates of the parameters of the chosen model, the calculated utilities of selected points in the attribute space, or a graph of the utility function. The decision maker may also view the implied indifference probabilities of gambles not used in the modelling phase.

If there are any modifications to the set of gamble triples and associated indifference probabilities, the modelling phase is entered again. This iteration continues until the utility function is acceptable to the decision maker.

Section 4. The Modelling Phase

The modelling phase is an iterative, non-linear least-squares algorithm. The objective function

(VII.4.1) $F = \Sigma[m(p_\theta) - m(\hat{p}_\theta)]^2$

is the sum of squared deviations of the assessed indifference probabilities from the implied indifference probabilities in the chosen metric. The algorithm attempts to find the minimum of equation (VII.4.1) with respect to the parameters of the chosen model for the utilities. There are three metrics and three models from which to chose.

The problem is to solve for the roots of the system of first-order derivatives of equation (VII.4.1). The Newton-Raphson method is used because it is stable and the necessary second-order derivatives are algebraically obtainable (although extremely complex). The new parameter estimates at each iteration are calculated by

(VII.4.2) $\underline{\eta}_{i+1} = \eta_i - H_i g_i$

where $\underline{\eta}_i$ is the vector of parameter estimates at iteration i, $\underline{g}_i$ is the gradient vector, the first derivatives, and H_i is the Hessian matrix, the inverse of the matrix of second-order derivatives. The gradient vector and the Hessian matrix of equation (VII.4.1) are presented in Appendix B.

The calculations of the Hessian matrix and the gradient vector are tedious and time-consuming. Therefore, a subiteration is also used. The correction of equation (VII.4.2) is applied until the minimum is overshot; this last correction is then rescinded. The correction is repeatedly halved, and applied (unless it overshoots the minimum), until the correction is negligible (on the order of 0.00001). The iteration continues with the recalculation of the gradient vector and the Hessian matrix. The iteration terminates when the sum of squared deviations differs from that of the previous iteration by a negligible amount (on the order of 0.00001).

Checks are made to insure that the parameter estimates remain within valid boundaries. If they exceed the boundaries, the correction is rescinded and halved, and the iteration continues. Note that the models insure that the utilities are monotone, so this need not be checked. To protect against a very slow convergence, the iterations are presented in sets of ten; the decision maker may terminate the estimation phase at the end of any set.

One additional problem may arise. If the utility estimates become too close, the numerical algorithm may become unstable. This may be observed by examining the derivatives used to calculate the improvements in equation

(VII.4.2). If the utility estimates used in any gamble triple to calculate the derivatives become essentially equal (e.g., within .03), that gamble triple is not used in that iteration of the modelling phase.

The metrics that are available are the indifference probability metric, the logodds metric, and the arcsine-square-root metric. The indifference probability metric implies the identity transformation:

(VII.4.3) $m(p) = p$.

The logodds metric

(VII.4.4) $m(p) = \log\left(\frac{p}{1-p}\right)$

and the arcsine-square-root metric

(VII.4.5) $m(p) = \arcsin(\sqrt{p})$

are widely used in Educational and Psychological research (Novick and Jackson, 1974) for transforming probabilities to obtain well-behaved quantities. The sensitivity of the solution to these metrics has been researched (Mayekawa, 1981) and found to be insignificant. They are made available for completeness. The necessary derivatives of these transformations are presented in Appendix B.

The three models that are available are three multivariate cumulative distribution functions: the multivariate normal, the Dirichlet, and the multivariate generalized beta. The normal and the Dirichlet distributions are well-known. Their univariate counterparts, the univariate normal and the standard beta distributions, have been used to model uniattribute utility functions (Novick and Lindley, 1980; Novick, Hamer, Libby, Chen, and Woodworth, 1980; Chen and Novick, 1981).

The generalized beta distribution was developed specifically for this application. The marginal distributions reproduce the same functional form, as do the conditional distributions with all but one dimension fixed. The univariate counterpart is a three-parameter beta-like distribution. The standard beta distribution can be shown to be a special case. The derivation of this distribution, as well as that of a companion distribution, are presented in Appendix A.

Because of the theory of conditional expected utility and the modelling of the utilities by cumulative distribution functions, we find the estimation may be greatly simplified. We need not work directly with the multivariate cumulative distribution function. By defining the conditional state sets as we have, we need only work with the corresponding conditional utility functions.

Each implied probability p_{ijk} in equation (VII.4.1) is calculated as a conditional utility, assuming that the three states in the gamble triple are all that matter. The calculation requires a utility for each state,

$$p_{ijk} = \frac{u(\theta_j) - u(\theta_i)}{u(\theta_k) - u(\theta_i)} \tag{VII.4.6}$$

but the utilities may be from any conditional utility function. They may be from the multiattribute utility function (every utility function is conditional). They may also be from the conditional utility function over the conditional set of states to which the gamble triple belongs. It is this latter approach that we take here.

Because of the manner in which we have defined our conditional set of states, these conditional utility functions are uniattribute. Because of the models that we have chosen to use, they are univariate cumulative distribution functions. The parameters of these conditional models are functions of the parameters of the multiattribute model. Thus, we may find the estimates of the parameters of the multiattribute utility function directly using this least-squares procedure. The necessary first- and second-order derivatives of the three models are presented in Appendix B.

Section 5. Discussion

The utility assessment procedure that has been proposed in this paper is based on sound theoretical and empirical judgments. The methods of elicitation, [illegible], and estimation have been proven effective through extensive research and use, although in a more restricted paradigm than proposed here. The theory of conditional expected utility justifies the application of the work in unidattribute utility assessment to the multiattribute paradigm.

The use of cumulative distribution functions greatly reduces the number of parameters in the model. While fewer parameters generally restrict the flexibility of the resulting fit, the models were selected with this in mind, and their use has been justified in the literature. A new probability density function, with great apparent flexibility, was developed specifically for this application.

The models that were proposed in this chapter have one significant disadvantage compared to the model proposed in Chapter V. Recall that in Chapter V, we proposed an extension of the Novick-Lindley utility assessment procedure based on results from multidimensional scaling and conjoint measurement. The advantage with that model is the potential for dimension reduction. The techniques are straightforward and were discussed in Chapter V. With the cumulative distribution function model, dimension reduction is possible, but not quite so easy.

The simplest form of dimension reduction for the cumulative distribution function utility model is marginalization. If the univariate marginal utility function in one dimension is essentially uniform, the dimension adds little to the overall multiattribute utility function. We could then marginalize out this dimension and have a simpler utility function.

This approach assumes that the attributes we have observed are the attributes with which we should be concerned. It may be the case that our interest lies along some subspace, such as a line, in the multidimensional attribute space. For example, in assessing utilities over Graduate Record Examination scores for a graduate school acceptance decision, it may be that we are concerned more with the sum of the verbal and quantitative scores than with the entire two-dimensional space.

The approach in this later case is considerably more complicated than simple marginalization. If we are concerned with linear transformations, the normal or t distribution would be a convenient form for investigating this type of dimension reduction because linear transformations reproduce the same functional form. We could determine other suitable distributions for known transformations, as well. The difficulty arises when we are not willing or able to specify the functional form of the transformation.

Some of the usual applications for the multiattribute utility assessment procedure may be suitable for the above type of analysis. Suppose that we have an attribute to be used in a decision. In a decision with risk, we will not know the value of this attribute when we must make the decision; we may, however, have a prediction equation for it using several available measures. For example, we may wish to decide to accept or reject a graduate school applicant based on first-year grades predicted by verbal and quantitative Graduate Record Examination scores.

If the prediction equation is linear, as they usually are, and if we are satisfied with a cumulative normal distribution function model for the utilities over the attribute of first-year grades, we should be able to determine a two-dimensional utility function over the GRE scores. This line of research is beyond the scope of this paper and is mentioned here only to indicate directions that it may take. It will be investigated in a separate study.

Appendix A The Generalized Beta and F Distributions

Section 1. Introduction

In developing the multiattribute utility function analysis based on cumulative distribution functions, it was necessary to derive a new probability density function of multiple dimensions. It may be that the functions presented here are two of the "many possible forms that are not in use at present and are not likely to have useful applications" (Johnson and Kotz, 1972).

In the following sections, two multivariate probability functions are derived. They are both based on generalizations of common distributions, namely the beta distribution and Snedecor's F distribution, and they are simple transformations of each other. The multivariate generalized beta distribution is discussed in Section 2 and the multivariate generalized F distribution is discussed in Section 3.

The moments of the generalized beta distribution are not computable in closed form. Approximations may be calculated from the moments of the generalized F distribution.

Both the approximate moments of the generalized beta distribution and the exact moments of the generalized F distribution are presented in Section 4.

Section 2. The Generalized Beta Distribution

The generalized beta distribution is a simple extension from the common beta distribution. It uses general gamma random variates instead of the chi-square random variates used in the derivation of the common beta distribution.

The impetus for the generalized beta distribution is as follows. For a model of a multiattribute utility function, we wanted a cumulative distribution function that had the same degree of flexibility in form that the beta distribution provided. Although the Dirichlet distribution is a multivariate generalization of the beta distribution, it has the restriction to a lower-dimensional simplex that is undesirable in some cases.

In particular, we wanted a multivariate distribution having positive probability on the full (0,1) hypercube, marginal distributions that reproduce the same distributional form, and "nice" conditional distributions. We were looking for a multivariate distribution whose univariate marginal distributions were beta distributions and which had some sort of standardized univariate conditional distributions that were beta distributions, much like we have in the Dirichlet distribution and the normal distribution.

We discovered a multivariate distribution with reproducing marginal distributions, univariate marginal distributions that are a generalization of the beta distribution, and univariate conditional distributions that are the same generalization of the beta distribution. We are calling this distribution a multivariate generalized beta distribution of the first kind.

Let $X_0, X_1, \ldots, X_r$ be distributed as independent gamma random variates with parameters α_i and β_i, $i = 0, \ldots, r$, respectively. The joint distribution of $X_0, X_1, \ldots, X_r$ is given by

$$\text{(A.2.1)} \qquad p(x_0, x_1, \ldots, x_r) = \prod_{i=0}^{r} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} x_i^{\alpha_i - 1} e^{-\beta_i x_i}, \quad 0 < x_i$$

Let $Y_0 = X_0$, and $Y_i = X_i/(1-X_i)$ for $i = 1, \ldots, r$. This gives the inverse transformation $X_0 = Y_0$, and $X_i = Y_0 Y_i/(1+Y_i)$ for $i = 1, \ldots, r$. The Jacobian for this distribution is given by

$$(A.2.2)\quad J[X\to Y] = \begin{vmatrix} 1 & \frac{y_1}{1-y_1} & \frac{y_2}{1-y_2} & \cdots & \frac{y_r}{1-y_r} \\ 0 & \frac{y_0}{(1-y_1)^2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \frac{y_0}{(1-y_r)^2} \end{vmatrix} = y_0^r \prod_{i=1}^{r} (1-y_i)^{-2}$$

Thus, the joint distribution of Y_0, Y_1, ..., Y_r is given by

$$(A.2.3)\quad p(y_0, y_1, \ldots, y_r) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)} y_0^{\alpha_0 - 1} e^{-\beta_0 y_0}$$

$$* \prod_{i=1}^{r} \left[\frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \left(\frac{y_0 y_i}{1-y_i} \right)^{\alpha_i - 1} e^{-\beta_i \left(\frac{y_0 y_i}{1-y_i} \right)} \right]$$

$$* \left[y_0^r \prod_{i=1}^{r} (1-y_i)^{-2} \right]$$

$$= \frac{\Gamma\left(\Sigma_{i=0}^{r} \alpha_i\right)}{\pi_{i=0}^{r} \Gamma(\alpha_i)} \; \frac{\beta_0^{\alpha_0} \prod_{i=1}^{r} \beta_i^{\alpha_i} \left(\frac{y_i}{1-y_i} \right)^{\alpha_i - 1} \left(\frac{1}{1-y_i} \right)^2}{\left[\beta_0 + \Sigma_{i=1}^{r} \beta_i \left(\frac{y_i}{1-y_i} \right) \right]^{\Sigma_{i=0}^{r} \alpha_i}}$$

$$* \partial\Gamma\left(\Sigma_{i=0}^{r} \alpha_i, \beta_0 + \Sigma_{i=1}^{r} \beta_i \left(\frac{y_i}{1-y_i} \right) \right)$$

where $\partial\Gamma\left(\sum_{i=0}^{r} \alpha_i, \beta_0 + \sum_{i=1}^{r} \beta_i \left(\frac{y_i}{1-y_i} \right) \right)$ represents the probability density function of a gamma-distributed random variate of parameters $\sum_{i=0}^{r} \alpha_i$ and $\beta_0 + \sum_{i=1}^{r} \beta_i \left(\frac{y_i}{1-y_i} \right)$. Thus the marginal distribution of Y_1, ..., Y_r is given by

(A.2.4) $$p(y_1,\ldots,y_r) = \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\prod_{i=0}^{r}\Gamma(\alpha_i)} \frac{\beta_0^{\alpha_0}\prod_{i=1}^{r}\left[\beta_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^{2}\right]}{\left[\beta_0+\sum_{i=1}^{r}\beta_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r}\alpha_i}}, \quad 0\le y_i<1$$

This may be rewritten to give

(A.2.5) $$p(y_1,\ldots,y_r) = \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\prod_{i=0}^{r}\Gamma(\alpha_i)} \frac{\prod_{i=1}^{r}\left[\lambda_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^{2}\right]}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r}\alpha_i}}, \quad 0\le y_i<1$$

where $\lambda_i = \beta_i/\beta_0$.

To find marginal distributions, noting that $\partial(Y_i/(1-Y_i))/\partial Y_i = (1-Y_i)^{-2}$ and integrating by parts, we find

(A.2.6) $$p(y_1,\ldots,y_{r-1}) = \int_0^1 p(y_1,\ldots,y_r)\,\partial y_r$$

$$= \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\prod_{i=0}^{r}\Gamma(\alpha_i)} \prod_{i=1}^{r-1}\left[\lambda_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^{2}\right]$$

$$* \int_0^1 \frac{\lambda_r^{\alpha_r}\left(\frac{y_r}{1-y_r}\right)^{\alpha_r-1}\left(\frac{1}{1-y_r}\right)^{2}\partial y_r}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r}\alpha_i}}$$

$$= \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\prod_{i=0}^{r}\Gamma(\alpha_i)} \frac{\prod_{i=1}^{r-1}\left[\lambda_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^2\right]}{\left[1+\sum_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r-1}\alpha_i}}$$

$$* \int_0^1 \frac{w_r^{\alpha_r - 1}\,\partial w_r}{\left[1+w_r\right]^{\sum_{i=0}^{r}\alpha_i}}$$

where

$$W_r = \frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{\left[1+\sum_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]} .$$

Therefore, the marginal distribution is given by

$$\text{(A.2.7)} \quad p(y_1,\ldots,y_{r-1}) = \frac{\Gamma\left(\sum_{i=0}^{r-1}\alpha_i\right)}{\prod_{i=0}^{r-1}\Gamma(\alpha_i)} \frac{\prod_{i=1}^{r-1}\left[\lambda_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^2\right]}{\left[1+\sum_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r-1}\alpha_i}}$$

since the distribution of W_r in equation (A.2.6) is a beta of the second kind and has a constant of proportionality $B^{-1}(\alpha_r, \sum_{i=0}^{r-1}\alpha_i)$.

From equation (A.2.7) we can see that the multivariate generalized beta distribution has the convenient property of reproduction of form: any subset of the variates again form a multivariate generalized beta distribution. Indeed, the univariate marginal distribution is given by

$$\text{(A.2.8)} \quad p(y_i) = \frac{\Gamma(\alpha_0+\alpha_i)}{\Gamma(\alpha_0)\Gamma(\alpha_i)} \frac{\lambda_i^{\alpha_i}\left(\frac{y_i}{1-y_i}\right)^{\alpha_i-1}\left(\frac{1}{1-y_i}\right)^2}{\left[1+\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\alpha_0+\alpha_i}} \qquad 0 < y_i < 1$$

thus reproducing the generalized beta form.

Note that if Y is distributed as a standard beta random variate of the first kind, with parameters α and β,

$$\text{(A.2.9)} \quad P(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1-y)^{\beta-1}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\left(\frac{y}{1-y}\right)^{\alpha-1}\left(\frac{1}{1-y}\right)^2}{\left[1+\left(\frac{y}{1-y}\right)\right]^{\alpha+\beta}}, \qquad 0 < y < 1$$

Hence, the standard beta distribution of the first kind is the generalized beta distribution with parameters α, β, and $\lambda = 1$.

Conditional distributions may be found in several ways. One may be found by observing that

$$(A.2.10)\quad p(y_r|y_1,\ldots,y_{r-1}) = \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\sum_{i=0}^{r-1}\alpha_i\right)} \; \frac{\lambda_r^{\alpha_r}\left(\frac{y_r}{1-y_r}\right)^{\alpha_r-1}\left(\frac{1}{1-y_r}\right)^{2}}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r}\alpha_i}}$$

$$* \left[1+\sum_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r-1}\alpha_i}$$

$$= \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\sum_{i=0}^{r-1}\alpha_i\right)} \; \frac{\lambda_r^{\alpha_r-1}\left(\frac{y_r}{1-y_r}\right)^{\alpha_r-1}}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\alpha_r-1}}$$

$$* \frac{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)-\lambda_r\left(\frac{y_r}{1-y_r}\right)\right]^{\sum_{i=0}^{r-1}\alpha_i}}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\sum_{i=0}^{r-1}\alpha_i}}$$

$$* \frac{\lambda_r\left(\frac{1}{1-y_r}\right)^{2}}{\left[1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]}$$

$$= \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\sum_{i=0}^{r-1}\alpha_i\right)} \left[\frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]^{\alpha_r-1}$$

$$* \left[1-\frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]^{\sum_{i=0}^{r-1}\alpha_i}$$

$$* \left[\frac{\lambda_r\left(\frac{1}{1-y_r}\right)^{2}}{1+\sum_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]$$

Let $y_r^{***} = \frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}$; then $\partial y_r^{***} = \left[1-\frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]$

$$* \left[\frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)^2 \partial y_r}{1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]$$

Thus

$$\text{(A.2.11)}\quad p(y_r^{***}|y_1,\ldots,y_{r-1}) = \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\Gamma\left(\Sigma_{i=0}^{r-1}\alpha_i\right)}(y_r^{***})^{\alpha_r-1}$$

$$*(1-y_r^{***})^{\Sigma_{i=0}^{r-1}\alpha_i-1}$$

$$= \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\Gamma\left(\Sigma_{i=0}^{r-1}\alpha_i\right)}$$

$$* \frac{\left(\frac{y_r^{***}}{1-y_r^{***}}\right)^{\alpha_r-1}\left(\frac{1}{1-y_r^{***}}\right)^2}{\left(1+\left(\frac{y_r^{***}}{1-y_r^{***}}\right)\right)^{\Sigma_{i=0}^{r}\alpha_i}} \qquad 0 \le y_r^{***} < \ldots$$

Hence, $Y_r^{***}|Y_1,\ldots,Y_{r-1}$ is distributed as a generalized beta random variate with parameters α_r, $\Sigma_{i=0}^{r-1}\alpha_i$, and $\lambda^{***} = 1$, i.e., a standard beta random variate of the first kind. Note that the moments of $Y_r^{***}|Y_1,\ldots,Y_{r-1}$ are not dependent on Y_1, ..., Y_{r-1}, and that the joint distribution of the r conditional random variates thus defined are constrained so that their sum is equal to one, as in the Dirichlet distribution.

An alternative conditional distributicn may be found by observing that

$$\text{(A.2.12)}\quad p(y_r|y_1,\ldots,y_{r-1}) = \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\Sigma_{i=0}^{r-1}\alpha_i\right)}\;\frac{\lambda_r^{\alpha_r}\left(\frac{y_r}{1-y_r}\right)^{\alpha_r-1}\left(\frac{1}{1-y_r}\right)^2}{\left[1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r}\alpha_i}}$$

$$* \left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r-1}\alpha_i}$$

$$= \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\Sigma_{i=1}^{r-1}\alpha_i\right)}\;\frac{\lambda_r^{\alpha_r-1}\left(\frac{y_i}{1-y_i}\right)^{\alpha_r-1}}{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\alpha_r-1}}$$

$$* \frac{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r}\alpha_i}}{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)+\lambda_r\left(\frac{y_r}{1-y_r}\right)\right]^{\Sigma_{i=0}^{r}\alpha_i}}$$

$$* \frac{\lambda_r\left(\frac{1}{1-y_r}\right)^2}{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]}$$

$$= \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\,\Gamma\left(\Sigma_{i=0}^{r-1}\alpha_i\right)}\left[\frac{\lambda_r\left(\frac{y_r}{1-y_r}\right)}{1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right]^{\alpha_r-1}$$

$$
* \left[1 + \frac{\lambda_r \left(\frac{y_r}{1-y_r} \right)}{1 + \Sigma_{i=1}^{r-1} \lambda_i \left(\frac{y_i}{1-y_i} \right)} \right]^{-\Sigma_{i=0}^{r} \alpha_i}
$$

$$
* \left[\frac{\lambda_r \left(\frac{1}{1-y_r} \right)^2}{1 + \Sigma_{i=1}^{r-1} \lambda_i \left(\frac{y_i}{1-y_i} \right)} \right]
$$

Let $y_r^{**} = \dfrac{\lambda_r \left(\frac{y_r}{1-y_r} \right)}{1 + \Sigma_{i=1}^{r-1} \lambda_i \left(\frac{y_i}{1-y_i} \right)}$; then $\partial y_r^{**} = \dfrac{\lambda_r \left(\frac{1}{1-y_r} \right)^2}{1 + \Sigma_{i=1}^{r-1} \lambda_i \left(\frac{y_i}{1-y_i} \right)} \partial y_r$, and

$$
\text{(A.2.13)} \quad p(y_r^{**} \mid y_1, \ldots, y_{r-1}) = \frac{\Gamma\left(\Sigma_{i=0}^{r} \alpha_i \right)}{\Gamma(\alpha_r) \, \Gamma\left(\Sigma_{i=0}^{r-1} \alpha_i \right)} \frac{\left(y_r^{**} \right)^{\alpha_r - 1}}{\left[1 + y_r^{**} \right]^{\Sigma_{i=0}^{r} \alpha_i}}, \quad 0 \le y_r^{**} .
$$

The distribution of Y_r^{**} is thus a beta distribution of the second kind. This is a special case of the distribution discussed in Section 3, the generalized F distribution, of this Appendix A. Note that $Y_r^{**} = Y_r^{***} / (1 - Y_r^{***})$, given $Y_1, \ldots, Y_{r-1}$; this relationship is developed further in Section 3 as well. Also note that the moments of Y_r^{**}, given $Y_1, \ldots, Y_{r-1}$, are not dependent on $Y_1, \ldots, Y_{r-1}$, but that the expected value of $Y_r/(1-Y_r)$ is linear in the odds ratios of the conditioning variates.

Another conditional distribution may be derived by observing that

$$\text{(A.2.14)}\quad p(y_r|y_1,\ldots,y_{r-1}) = \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\Gamma\left(\Sigma_{i=0}^{r-1}\alpha_i\right)}$$

$$* \frac{\lambda_r^{\alpha_r}\left(\frac{y_r}{1-y_r}\right)^2}{\left[1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r}\alpha_i}}$$

$$* \left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r-1}\alpha_i}$$

$$= \frac{\Gamma\left(\Sigma_{i=0}^{r}\alpha_i\right)}{\Gamma(\alpha_r)\Gamma\left(\Sigma_{i=1}^{r-1}\alpha_i\right)}$$

$$* \frac{\lambda_r^{\alpha_r}\left(\frac{y_r}{1-y_r}\right)^{\alpha_r-1}\left(\frac{1}{1-y_r}\right)^2}{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\alpha_r}}$$

$$* \frac{\left[1+\Sigma_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r-1}\alpha_i}}{\left[1+\Sigma_{i=1}^{r}\lambda_i\left(\frac{y_i}{1-y_i}\right)\right]^{\Sigma_{i=0}^{r-1}\alpha_i}}$$

$$[illegible]$$

$$\cdot \frac{\left(\frac{y_r}{1-y_r}\right)^{\alpha_r - 1} \Big/ (1-y_r)^2}{\left[1+\left(\frac{\lambda_r}{1+\sum_{i=1}^{r-1}\lambda_i\left(\frac{y_i}{1-y_i}\right)}\right)\left(\frac{y_r}{1-y_r}\right)\right]^{\sum_{i=0}^{r}\alpha_i}}$$

Hence, $Y_r^* = Y_r \mid Y_1, \ldots, Y_{r-1}$ is distributed as a generalized beta random variate with parameters $\alpha^* = \alpha_r$, $\beta^* = \sum_{i=0}^{r-1} \alpha_i$, and

$$\lambda^* = \frac{\lambda_r}{1 + \sum_{i=0}^{r-1} \lambda_i \left(\frac{y_i}{1-y_i}\right)}.$$

Except for the result that the moments of the generalized beta distribution are not obtainable in closed form, this conditional distribution is the most useful of the three. It relates directly to Y_r, not to the odds ratio of Y_r. It is the same distributional form as the multivariate and marginal distributions. Finally, the moments are likely to be dependent on $Y_1, \ldots, Y_r$, as odds ratios, through the parameter λ^*.

The fact that the moments are not obtainable in closed form is a drawback to its usefulness. We may, however, obtain approximations to the moments, using the companion distribution discussed in Section 3. The moments of both distributions are discussed further in Section 4.

Section 3. The Generalized F Distribution

The generalized F distribution is a simple generalization of the standard, Snedecor's F distribution. It may be derived from an underlying gamma distribution like the standard F distribution is derived from the chi-square distribution. Note that the chi-square distribution is a special case of the gamma distribution. The generalized F distribution may also be derived as a transformation of the generalized beta distribution, in much the same way as the standard F distribution may be derived as a transformation from the standard beta distribution. In fact, it was derived initially this way by this author.

Both derivations are presented below, with the derivation from the gamma random variates first. As in the previous section, the development is in terms of the multivariate distribution.

Let X_0, X_1, ..., X_r be distributed as independent gamma random variates with parameters [illegible] and [illegible], $i=0, 1, \ldots, r$ respectively. The joint distribution of X_0, X_1, ..., X_r is given by

$$\text{(A.3.1)}\quad p(x_0, x_1, \ldots, x_r) = \prod_{i=0}^{r}\left[\frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)}\, x_i^{\alpha_i - 1}\, e^{-\beta_i x_i}\right],\quad 0 \le x_i .$$

Let $Z_0 = X_0$, and $Z_i = X_i/X_0$ for i=1, ..., r. This gives the inverse transformaton $X_0 = Z_0$, and $X_i = Z_0 * Z_i$ for i=1, ..., r. The Jacobian for this transformation is given by

$$\text{(A.3.2)}\quad J[X \to Z] = \begin{vmatrix} 1 & z_1 & z_2 & \cdots & z_r \\ 0 & z_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & z_0 \end{vmatrix} = z_0^{r}$$

Thus, the joint distribution of Z , Z , ..., Z is given by

$$\text{(A.3.3)}\quad p(z_0, z_1, \ldots, z_r) = \frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\, z_0^{\alpha_0 - 1}\, e^{-\beta_0 z_0}$$

$$* \prod_{i=1}^{r}\left[\frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)}(z_0 z_i)^{\alpha_i - 1}\, e^{-\beta_i z_0 z_i}\right] z_0^{r}$$

$$= \frac{\Gamma\left(\sum_{i=0}^{r}\alpha_i\right)\ \beta_0^{\alpha_0}\prod_{i=0}^{r}\left[\beta_i^{\alpha_i}\ z_i^{\alpha_i - 1}\right]}{\prod_{i=0}^{r}\Gamma(\alpha_i)\left[\beta_0 + \sum_{i=1}^{r}\beta_i z_i\right]^{\sum_{i=0}^{r}\alpha_i}}$$

$$* \ \mathrm{Gr}\left(\sum_{i=0}^{r}\alpha_i,\ \beta_0 + \sum_{i=1}^{r}\beta_i z_i\right)$$

where $\mathrm{Gr}\left(\sum_{i=0}^{r}\alpha_i,\ \beta_0 + \sum_{i=1}^{r}\beta_i z_i\right)$ represents the probability density function of a gamma-distributed random variate with parameters $\sum_{i=0}^{r} \alpha_i$ and $\beta_0 + \sum_{i=1}^{r} \beta_i z_i$. Thus the marginal distribution of $Z_1, \ldots, Z_r$ is given by

$$(A.3.4) \quad p(z_1, \ldots, z_r) = \frac{\Gamma\left(\sum_{i=0}^{r} \alpha_i\right)}{\prod_{i=0}^{r} \Gamma(\alpha_i)} \; \frac{\prod_{i=1}^{r} \lambda_i^{\alpha_i} z_i^{\alpha_i - 1}}{\left[1 + \sum_{i=1}^{r} \lambda_i z_i\right]^{\sum_{i=0}^{r} \alpha_i}}, \quad 0 \le z_i$$

where $\lambda_i = \beta_i / \beta_0$, as in Section 2. This is the multivariate generalized F distribution.

Marginal distributions may be found by integration, so that

$$(A.3.5) \quad p(z_1, \ldots, z_{r-1}) = \frac{\Gamma\left(\sum_{i=0}^{r-1} \alpha_i\right)}{\prod_{i=0}^{r-1} \Gamma(\alpha_i)} \; \frac{\prod_{i=1}^{r-1} \lambda_i^{\alpha_i} z_i^{\alpha_i - 1}}{\left[1 + \sum_{i=1}^{r-1} \lambda_i z_i\right]^{\sum_{i=0}^{r-1} \alpha_i}}, \quad 0 \le z_i$$

and

$$(A.3.6) \quad p(z_i) = \frac{\Gamma(\alpha_0 + \alpha_i)}{\Gamma(\alpha_0)\Gamma(\alpha_i)} \; \frac{\lambda_i^{\alpha_i} z_i^{\alpha_i - 1}}{\left[1 + \lambda_i z_i\right]^{\alpha_0 + \alpha_i}}, \quad 0 \le z_i$$

Thus, the multivariate generalized F distribution has the convenient reproduction of distributional form in any subset marginalization, including the univariate marginal distributions.

Note that the standard, Snedecor's F distribution is a special case of this generalized F distribution. If Z is distributed as a standard F random variate, its probability density function may be written

$$\text{(A.3.7)} \quad p(z) = \frac{\Gamma\left(\frac{1}{2}k_1+\frac{1}{2}k_2\right)}{\Gamma\left(\frac{1}{2}k_1\right)\Gamma\left(\frac{1}{2}k_2\right)} \; \frac{\left(k_1/k_2\right)^{\frac{1}{2}k_1} z^{\frac{1}{2}k_1-1}}{\left[1+(k_1/k_2)z\right]^{\frac{1}{2}k_1+\frac{1}{2}k_2}} \qquad 0 \le z .$$

Thus, it is a generalized F random variate with parameters $\frac{1}{2}k_1$, $\frac{1}{2}k_2$, and $\lambda = k_1/k_2$.

Note also that the beta distribution of the second kind is a special case of the generalized F distribution, with $\lambda = 1$. In addition, the multivariate F distribution and the multivariate inverted beta distribution mentioned by Johnson and Kotz (1972) are special cases of the multivariate generalized F distribution in the same way as are the standard F distribution and the beta distribution of the second kind of the univariate generalized F distribution.

The generalized F distribution may also be derived from the generalized beta distribution as follows. Let Y be distributed as a generalized beta random variate with parameters α, β, and λ. Let $Z = Y/(1-Y)$; then $\partial Z = (1-Y)^{-2}\,\partial Y$ and

$$(A.3.8) \quad p(z) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\lambda^{\alpha} z^{\alpha-1}}{[1 + \lambda_z]^{\alpha+\beta}}, \quad 0 \le z,$$

which is the probability density function of a generalized F random variate with parameters α, β, and λ. The same transformation applies in the multivariate case as may be seen by comparing (A.2.5) and (A.3.4).

Conditional distributions may be found in ways similar to those presented in Section 2. If we define

$$(A.3.9) \quad z_r^{***} = \frac{\lambda_r z_r}{\left[1 + \Sigma_{i=1}^{r} \lambda_i z_i\right]},$$

we find that Z_r^{***} is identical in distribution to Y_r^{***} of Section 2. Similarly, if we define

$$(A.3.10) \quad z_r^{**} = \frac{\lambda_r z_r}{1 + \Sigma_{i=1}^{r-1} \lambda_i z_i}$$

we find that Z_r^{**} is identical in distribution to Y_r^{**}.

Lastly, we may find may the distribution of Z_r^{*} = $Z_r | Z_1, \ldots, Z_{r-1}$ is that of a generalized F random variate with parameter α_r, $\Sigma_{i=1}^{r-1} \alpha_i$, and

$$(A.3.11) \quad \lambda^{*} = \frac{\lambda_r}{1 + \Sigma_{i=1}^{r-1} \lambda_i z_i}.$$

Note that Z_r^* is a transformation of Y_r^* of Section 2 such that $Z_r^* = Y_r^* / (1-Y_r^*)$.

The moments of the generalized F distribution are algebraically calculable, and are presented in the next section.

Section 4. The Moments of the Distributions

In this section we discuss the moments of the generalized beta distribution and of the generalized F distribution. The moments of the latter are algebraically calculable, whereas those of the former are not. We will derive the algebraic forms of the moments for the generalized F distribution first, and then derive approximations for the moments of the generalized beta distribution from them using Taylor's series expansion of the transformation from the generalized F distribution to the generalized beta distribution.

Let Z be distributed as a generalized F random variate with parameters α, β, and λ:

$$(A.4.1)\quad p(z) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\lambda^{\alpha} z^{\alpha-1}}{[1+\lambda z]^{\alpha+\beta}}, \qquad 0 \leq z$$

The t-th moment about the origin is given by

$$\text{(A.4.2)} \quad \xi[z^t] = \int_0^\infty \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\lambda^\alpha z^{\alpha+t-1}}{[1+\lambda z]^{\alpha+\beta}} \partial z$$

$$= \frac{\Gamma(\alpha+t)\Gamma(\beta-t)}{\Gamma(\alpha)\Gamma(\beta)\lambda^t}, \qquad t < \beta.$$

For $t \geq \beta$, the moments are infinite. Thus, the mean μ_z, variance σ_z^2, and skewness γ_z of this distribution are given by

$$\text{(A.4.3)} \quad \mu_z = \frac{\alpha}{(\beta-1)\lambda}, \qquad \beta > 1$$

$$\sigma_z^2 = \frac{\alpha(\alpha+\beta-1)}{(\beta-1)(\beta-2)\lambda^2}, \qquad \beta > 2$$

$$\gamma_z = \frac{\alpha(6\alpha(\beta-1)+\beta(\beta-2)+1)}{(\beta-1)^3(\beta-2)(\beta-1)\lambda^3}, \qquad \beta > 3.$$

Other central moments may be found in a similar manner, by expanding to a function of the moments about the origin and using (A.4.2). These three are all that are used in the approximations below.

Since the moments of the generalized beta distribution are not algebraically calculable, we now present some approximations. The approximations are based on the transformations from the generalized F distribution to the generalized beta distribution and Taylor's series expansion of it.

Suppose we have a random variate Z that has a generalized F distribution. Then the moments of a random variate $Y = \phi(Z)$ may be found if the transformation $\phi(\cdot)$ has derivatives to all degrees at least in a interval about some point used in the expansion, in this case the mean of Z, μ_z.

By Taylor's series expansion,

$$\text{(A.4.4)} \quad y = \phi[z] \doteq \phi[\mu_z] + \phi'[\mu_z](z-\mu_z) + \frac{1}{2}\phi''[\mu_z](z-\mu_z)^2 + \frac{1}{6}\phi'''[\mu_z](z-\mu_z)^3$$

Hence, the mean of Y is given by, approximately,

$$\text{(A.4.5)} \quad E[y] \doteq \phi[\mu_z] + \frac{1}{2}\phi''[\mu_z]\sigma_z^2 + \frac{1}{6}\phi'''[\mu_z]\gamma_z$$

Similarly,

$$\text{(A.4.6)} \quad V[y] \doteq (\phi'[\mu_z])^2\sigma_z^2 + \phi'[\mu_z]\phi''[\mu_z]\gamma_z - \frac{1}{4}(\phi''[\mu_z])^2\sigma_z^4 - \frac{1}{6}\phi''[\mu_z]\phi'''[\mu_z]\sigma_z^2\gamma_z - \frac{1}{36}(\phi'''[\mu_z])^2\gamma_z^2$$

$$\text{(A.4.7)} \quad E[(y-E[y])^3] \doteq (\phi'[\mu_z])^3\gamma_z + \frac{1}{4}(\phi''[\mu_z])^3\sigma_z^6 + \frac{1}{108}(\phi'''[\mu_z])^3\gamma_z^3 - \frac{3}{2}(\phi'[\mu_z])^2\phi''[\mu_z]\sigma_z^4$$

$$- \quad \frac{1}{2}(\phi'[\mu_z])^2\phi'''[\mu_z]\sigma_z^{\,2}\gamma_z$$

$$- \frac{3}{2}\phi'[\mu_z](\phi''[\mu_z])^2\sigma_z^{\,2}\gamma_z$$

$$+ \quad \frac{1}{4}(\phi''[\mu_z])^2\phi'''[\mu_z]\sigma_z^{\,4}\gamma_z$$

$$+ \frac{1}{12}\phi''[\mu_z](\phi''[\mu_z])^2\sigma_z^{\,2}\gamma_z^{\,2}$$

$$- \quad \frac{1}{2}\phi'[\mu_z]\phi''[\mu_z]\phi'''[\mu_z]\gamma_z^{\,2}.$$

In the above three equations,

$$\text{(A.4.8)} \qquad \phi[z] = \frac{z}{1 + z}$$

$$\phi'[z] = (1 + z)^{-2}$$

$$\phi''[z] = -2(1 + z)^{-3}$$

$$\phi'''[z] = 6(1 + z)^{-4}$$

and the moments of the random variate Z are taken from equations (A.4.3).

Section 5. Conclusion

In this appendix, we have presented two probability distributions that have proven useful in utility modelling. Both multivariate and univariate forms have been developed. The convenient property of reproduction of distributional form has been demonstrated for marginal distributions. Several conditional distributions have been derived.

It has been shown that these two distributions, called the generalized beta distribution and the generalized F distribution, reduce to the standard beta and F distributions, respectively, in special cases. It has also been shown that the two generalized distributions have a relationship similar to that between the standard beta and F distributions. Finally, the moments of the generalized F distribution have been presented, and approximations for the first three central moments of the generalized beta distribution have been derived.

Appendix B. Derivatives

Section 1. Introduction

In this appendix, we present the derivatives of the three probability distributions used in this paper to model multiattribute utility functions. Thse derivatives are necessary for the non-linear least-squares algorithm used to estimate the parameters of the model selected. The models that are available are the cumulative distribution functions of the multivariate normal, the Dirichlet, and the multivariate generalized beta distributions. Since only the univariate conditional distributions are required for the estimation procedure, the derivatives presented here are of the relevant cumulative univariate conditional distribution functions.

Because the derivatives are very complex, and because we are implementing three models of utilities in three metrics, the derivatives are presented in pieces. In Section 2, we present the derivatives of the objective function, that which we strive to minimize, with respect to an arbitrary estimate of the indifference probability. As we are implementing three metrics in the model estimation phase, the derivatives of these transformations are relevant as well. Thus, the derivatives of the objective function are presented for each of the three metric: the indifference probability metric, the logodds metric, and the arcsine-square-root metric. The estimate of the indifference probability is a function of the conditional utility function. The derivatives of the estimate of the indifference probability with respect to an arbitrary utility function is also presented here.

The conditional utility functions are modelled by three cumulative probability distribution functions: the normal, the beta, and the generalized beta. In Sections 3, 4, and 5, the derivatives of the utility function with respect to its parameters are presented for the three models, respectively. Each section is divided into two parts: the first part gives the derivatives of the utility function with respect to the conditional parameters; and the second part gives the derivatives of the conditional parameters with respect to the unconditional parameters (the parameters of the multivariate distribution that is the model of the multiattribute utility function).

Section 2. Estimation Metric Derivatives

Indifference Probability Metric:

$$F_I(\hat{p}) = \sum_{ijk} [p_{ijk} - \hat{p}_{ijk}]^2$$

$$\frac{\partial}{\partial \tau_s} F_I(\hat{p}) = -2 \sum_{ijk} [p_{ijk} - \hat{p}_{ijk}] \left(\frac{\partial}{\partial \tau_s} \hat{p}_{ijk}\right)$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t} F_I(\hat{p}) = 2 \sum_{ijk} \left\{ \left(\frac{\partial}{\partial \tau_s} \hat{p}_{ijk}\right) \left(\frac{\partial}{\partial \tau_t} \hat{p}_{ijk}\right) - [p_{ijk} - \hat{p}_{ijk}] \left(\frac{\partial^2}{\partial \tau_s \partial \tau_t} \hat{p}_{ijk}\right) \right\}$$

Logodds Metric:

$$F_L(\hat{p}) = \sum_{ijk} \left[\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) - \log\left(\frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}}\right)\right]^2$$

$$\frac{\partial}{\partial \tau_s} F_L(\hat{p}) = -2 \sum_{ijk} \left[\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) - \log\left(\frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}}\right)\right] * \left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)\left(\frac{\partial}{\partial \tau_s} \hat{p}_{ijk}\right)$$

$$\begin{aligned}\frac{\partial^2}{\partial \tau_s \partial \tau_t} F_L(\hat{p}) = 2 \sum_{ijk} \Bigg\{ & \left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)^2 \left(\frac{\partial}{\partial \tau_s} \hat{p}_{ijk}\right)\left(\frac{\partial}{\partial \tau_t} \hat{p}_{ijk}\right) \\ & - \left[\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) - \log\left(\frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}}\right)\right]\left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right) \\ & * \left(\frac{\partial^2}{\partial \tau_s \partial \tau_t} \hat{p}_{ijk}\right) \\ & + \left[\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) - \log\left(\frac{\hat{p}_{ijk}}{1-\hat{p}_{ijk}}\right)\right]\left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right) \\ & * \left(\frac{1-2\hat{p}_{ijk}}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)\left(\frac{\partial}{\partial \tau_s} \hat{p}_{ijk}\right)\left(\frac{\partial}{\partial \tau_t} \hat{p}_{ijk}\right)\end{aligned}$$

Arcsine Square Root Metric:

$$F_A(\hat{p}) = \sum_{ijk}\left[\arcsin\left(\sqrt{p_{ijk}}\right) - \arcsin\left(\sqrt{\hat{p}_{ijk}}\right)\right]^2$$

$$\frac{\partial}{\partial\tau_s}F_A(\hat{p}) = -\sum_{ijk}\left[\arcsin\left(\sqrt{p_{ijk}}\right) - \arcsin\left(\sqrt{\hat{p}_{ijk}}\right)\right] * \left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)^{\frac{1}{2}}\left(\frac{\partial}{\partial\tau_s}\hat{p}_{ijk}\right)$$

$$\begin{aligned}
\frac{\partial^2}{\partial\tau_s\partial\tau_t}F_A(\hat{p}) = \sum_{ijk}\Bigg\{&\left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)\left(\frac{\partial}{\partial\tau_s}\hat{p}_{ijk}\right)\left(\frac{\partial}{\partial\tau_t}\hat{p}_{ijk}\right)\\
&- \left[\arcsin\left(\sqrt{p_{ijk}}\right) - \arcsin\left(\sqrt{\hat{p}_{ijk}}\right)\right]\\
&* \left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)^{\frac{1}{2}}\left(\frac{\partial^2}{\partial\tau_s\partial\tau_t}\hat{p}_{ijk}\right)\\
&+\frac{1}{2}\left[\arcsin\left(\sqrt{p_{ijk}}\right) - \arcsin\left(\sqrt{\hat{p}_{ijk}}\right)\right]\\
&* \left(\frac{1}{\hat{p}_{ijk}(1-\hat{p}_{ijk})}\right)^{\frac{3}{2}}(1-2\hat{p}_{ijk})\left(\frac{\partial}{\partial\tau_s}\hat{p}_{ijk}\right)\\
&* \left(\frac{\partial}{\partial\tau_t}\hat{p}_{ijk}\right)\Bigg\}
\end{aligned}$$

Indifference Probability Estimate:

$$\hat{p}_{ijk} = \frac{\hat{u}(\theta_j) - \hat{u}(\theta_i)}{\hat{u}(\theta_k) - \hat{u}(\theta_i)}$$

$$\frac{\partial}{\partial \tau_s}\hat{p}_{ijk} = \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)^2 \left[\left(\hat{u}(\theta_k)-\hat{u}(\theta_i)\right)\left(\frac{\partial}{\partial \tau_s}\hat{u}(\theta_j) - \frac{\partial}{\partial \tau_s}\hat{u}(\theta_i)\right)\right.$$

$$\left. - \left(\frac{\partial}{\partial \tau_s}\hat{u}(\theta_k) - \frac{\partial}{\partial \tau_s}\hat{u}(\theta_i)\right)\left(\hat{u}(\theta_j)-\hat{u}(\theta_i)\right)\right]$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t}\hat{p}_{ijk} = \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)\left(\frac{\partial^2}{\partial \tau_s \partial \tau_t}\hat{u}(\theta_j) - \frac{\partial^2}{\partial \tau_s \partial \tau_t}\hat{u}(\theta_i)\right)$$

$$- \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)^2 \left(\frac{\partial^2}{\partial \tau_s \partial \tau_t}\hat{u}(\theta_k) - \frac{\partial^2}{\partial \tau_s \partial \tau_t}\hat{u}(\theta_i)\right)$$

$$* \left(\hat{u}(\theta_j) - \hat{u}(\theta_i)\right)$$

$$- \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)^2 \left(\frac{\partial}{\partial \tau_s}\hat{u}(\theta_k) - \frac{\partial}{\partial \tau_s}\hat{u}(\theta_i)\right)$$

$$* \left(\frac{\partial}{\partial \tau_t}\hat{u}(\theta_j) - \frac{\partial}{\partial \tau_t}\hat{u}(\theta_i)\right)$$

$$- \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)^2 \left(\frac{\partial}{\partial \tau_t}\hat{u}(\theta_k) - \left(\frac{\partial}{\partial \tau_t}\hat{u}(\theta_i)\right)\right.$$

$$* \left(\frac{\partial}{\partial \tau_s}\hat{u}(\theta_j) - \frac{\partial}{\partial \tau_s}\hat{u}(\theta_i)\right)$$

$$+2 \left(\frac{1}{\hat{u}(\theta_k)-\hat{u}(\theta_i)}\right)^3 \left(\frac{\partial}{\partial \tau_t}\hat{u}(\theta_k) - \frac{\partial}{\partial \tau_t}\hat{u}(\theta_i)\right)$$

$$* \left(\frac{\partial}{\partial \tau_s}\hat{u}(\theta_k) - \frac{\partial}{\partial \tau_s}\hat{u}(\theta_i)\right)\left(\hat{u}(\theta_j) - \hat{u}(\theta_i)\right)$$

Section 3. Normal Distribution

$$u(\theta) = \Phi\left(\frac{\theta-\mu}{\sigma}\right) = \int_{-\infty}^{\frac{\theta-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

$$\frac{\partial}{\partial\tau_s}u(\theta) = \phi\left(\frac{\theta-\mu}{\sigma}\right)\left[\frac{1}{\sigma^2}\left(\sigma\frac{\partial}{\partial\tau_s}\mu + (\theta-\mu)\frac{\partial}{\partial\tau_s}\sigma\right)\right]$$

$$\frac{\partial^2}{\partial\tau_s\partial\tau_t}u(\theta) = \phi\left(\frac{\theta-\mu}{\sigma}\right)\left[\frac{1}{\sigma^2}\left(\sigma\frac{\partial}{\partial\tau_s}\mu + (\theta-\mu)\frac{\partial}{\partial\tau_s}\sigma\right)\right]$$

$$\times\left[\frac{1}{\sigma^2}\left(\sigma\frac{\partial}{\partial\tau_t}\mu + (\theta-\mu)\frac{\partial}{\partial\tau_t}\sigma\right)\right]\left[\left(\frac{\theta-\mu}{\sigma}\right)\right]$$

$$+\phi\left(\frac{\theta-\mu}{\sigma}\right)\left[\frac{1}{\sigma}\left(\frac{\partial^2}{\partial\tau_s\partial\tau_t}\mu\right) - \frac{1}{\sigma^2}\left(\frac{\partial}{\partial\tau_s}\mu\right)\left(\frac{\partial}{\partial\tau_t}\sigma\right)\right.$$

$$-\frac{1}{\sigma^2}\left(\frac{\partial}{\partial\tau_s}\sigma\right)\left(\frac{\partial}{\partial\tau_t}\mu\right) - \frac{2(\theta-\mu)}{\sigma^3}\left(\frac{\partial}{\partial\tau_s}\sigma\frac{\partial}{\partial\tau_t}\sigma\right)$$

$$\left.+\frac{(\theta-\mu)}{\sigma^2}\left(\frac{\partial^2}{\partial\tau_s\partial\tau_t}\sigma\right)\right]\left(\frac{\theta-\mu}{\sigma}\right)$$

Marginal Parameters:

$$\mu = \mu_i, \qquad \sigma = \sigma_i$$

$$\frac{\partial}{\partial\tau_s}\mu = \begin{cases}1 & \tau_s = \mu_i\\ 0 & \text{otherwise}\end{cases} \qquad \frac{\partial}{\partial\tau_s}\sigma = \begin{cases}1 & \tau_s = \sigma_i\\ 0 & \text{otherwise}\end{cases}$$

$$\frac{\partial^2}{\partial\tau_s\partial\tau_t}\mu = \frac{\partial^2}{\partial\tau_s\partial\tau_t}\sigma = 0$$

Conditional Parameters:

$$\mu = \mu_i + \Sigma_{ij}\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j) \qquad \sigma = \left(\Sigma_{ii} - \Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}\right)^{\frac{1}{2}}$$

$$\frac{\partial}{\partial \tau_s}\mu = \frac{\partial}{\partial \tau_s}\mu_i + \left(\frac{\partial}{\partial \tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j) - \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j)$$

$$- \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\underline{\mu}_j\right)$$

$$\frac{\partial}{\partial \tau_s}\sigma = \frac{1}{2}\left(\Sigma_{ii} - \Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}\right)^{-\frac{1}{2}}$$

$$\star\left[\left(\frac{\partial}{\partial \tau_s}\Sigma_{ii}\right) - \left(\frac{\partial}{\partial \tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\Sigma_{ji} + \Sigma_{ij}\Sigma_{jj}^{-1}\right.$$

$$\left.\star\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji} - \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{ji}\right)\right]$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t}\mu = -\left(\frac{\partial}{\partial \tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j)$$

$$- \left(\frac{\partial}{\partial \tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\underline{\mu}_j\right)$$

$$- \left(\frac{\partial}{\partial \tau_t}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j)$$

$$+ \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j)$$

$$+ \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}(\underline{\theta}_j - \underline{\mu}_j)$$

$$+ \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\underline{\mu}_j\right)$$

$$- \left(\frac{\partial}{\partial \tau_t}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\underline{\mu}_j\right)$$

$$+ \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial \tau_s}\underline{\mu}_j\right)$$

$$
\begin{aligned}
\frac{\partial^2}{\partial\tau_s\partial\tau_t}\rho = & -\frac{1}{4}\left(\Sigma_{ii}-\Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}\right)^{-\frac{3}{2}} \\
& * \left[\left(\frac{\partial}{\partial\tau_s}\Sigma_{ii}\right) - \left(\frac{\partial}{\partial\tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\Sigma_{ji}+\Sigma_{ij}\Sigma_{jj}^{-1}\right. \\
& * \left.\left(\frac{\partial}{\partial\tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji}-\Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{ji}\right)\right] \\
& * \left[\left(\frac{\partial}{\partial\tau_t}\Sigma_{ii}\right) - \left(\frac{\partial}{\partial\tau_t}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\Sigma_{ji}+\Sigma_{ij}\Sigma_{jj}^{-1}\right. \\
& * \left.\left(\frac{\partial}{\partial\tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji}-\Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{ji}\right)\right] \\
& +\frac{1}{2}\left(\Sigma_{ii}-\Sigma_{ij}\Sigma_{jj}^{-1}\Sigma_{ji}\right)^{-\frac{1}{2}} \\
& * \left[\left(\frac{\partial}{\partial\tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji}\right. \\
& -\left(\frac{\partial}{\partial\tau_s}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{ji}\right) \\
& +\left(\frac{\partial}{\partial\tau_t}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji} \\
& - \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji} \\
& - \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\Sigma_{ji} \\
& + \Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{ji}\right) \\
& - \left(\frac{\partial}{\partial\tau_t}\Sigma_{ij}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{ji}\right) \\
& + \left.\Sigma_{ij}\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_t}\Sigma_{jj}\right)\Sigma_{jj}^{-1}\left(\frac{\partial}{\partial\tau_s}\Sigma_{ji}\right)\right]
\end{aligned}
$$

Assuming a two-attribute problem, we may simplify these equations. Suppose that the parameter vector is $\underline{\theta}' = (\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22})$, where μ_1 and μ_2 are means, σ_{11} and σ_{22} are variances, and σ_{12} is the covariance. Then

$$\frac{\partial}{\partial \tau_s}\mu = \begin{cases} 1 & \tau_s = \mu_1 \\ -\sigma_{12}/\sigma_{22} & \tau_s = \mu_2 \\ 0 & \tau_s = \sigma_{11} \\ (x_2-\mu_2)/\sigma_{22} & \tau_s = \sigma_{12} \\ -\sigma_{12}(x_2-\mu_2)/\sigma_{22} & \tau_s = \sigma_{22} \end{cases}$$

$$\frac{\partial}{\partial \tau_s}\sigma = \begin{cases} 0 & \tau_s = \mu_1 \\ 0 & \tau_s = \mu_2 \\ \frac{1}{2}\left(\sigma_{11} - \frac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{1}{2}} & \tau_s = \sigma_{11} \\ -\left(\frac{\sigma_{12}}{\sigma_{22}}\right)\left(\sigma_{11} - \frac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{1}{2}} & \tau_s = \sigma_{12} \\ \frac{1}{2}\left(\frac{\sigma_{12}}{\sigma_{22}}\right)^{2}\left(\sigma_{11} - \frac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{1}{2}} & \tau_s = \sigma_{22} \end{cases}$$

$$
\frac{\partial^2}{\partial\tau_s\,\partial\tau_t}\,\mu=\begin{cases}
-\dfrac{1}{\sigma_{22}} & \tau_s=\mu_2,\ \tau_t=\sigma_{12}\\[2ex]
\dfrac{\sigma_{12}}{\sigma_{22}^{2}} & \tau_s=\mu_2,\ \tau_t=\sigma_{22}\\[2ex]
-\dfrac{\left(\theta_2-\mu_2\right)}{\sigma_{22}^{2}} & \tau_s=\sigma_{12},\ \tau_t=\sigma_{22}\\[2ex]
2\,\sigma_{12}(\theta_2-\mu_2)/\sigma_{22}^{3} & \tau_s=\tau_t=\sigma_{22}\\[2ex]
0 & \text{otherwise}
\end{cases}
$$

$$
\frac{\partial^2}{\partial\tau_s\,\partial\tau_t}\,\sigma=\begin{cases}
-\dfrac{1}{4}\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\tau_t=\sigma_{11}\\[2ex]
\dfrac{1}{2}\left(\dfrac{\sigma_{12}}{\sigma_{22}}\right)\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\sigma_{11},\ \tau_t=\sigma_{12}\\[2ex]
-\dfrac{1}{4}\left(\dfrac{\sigma_{12}}{\sigma_{22}}\right)^{2}\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\sigma_{11},\ \tau_t=\sigma_{22}\\[2ex]
-\left(\sigma_{11}-\left(\dfrac{\sigma_{12}}{\sigma_{22}}\right)^{2}\left(\sigma_{22}^{-1}\right)\right)\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\tau_t=\sigma_{12}\\[2ex]
\dfrac{1}{2}\left(\dfrac{\sigma_{12}}{\sigma_{22}^{3}}\right)\left(2\sigma_{11}\sigma_{22}-\sigma_{12}^{2}\right)\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\sigma_{12},\ \tau_t=\sigma_{22}\\[2ex]
-\dfrac{1}{4}\left(\dfrac{\sigma_{12}^{2}}{\sigma_{22}^{4}}\right)\left(4\sigma_{11}\sigma_{22}-3\sigma_{12}^{2}\right)\left(\sigma_{11}-\dfrac{\sigma_{12}^{2}}{\sigma_{22}}\right)^{-\frac{3}{2}} & \tau_s=\tau_t=\sigma_{22}
\end{cases}
$$

Section 4. Standard Beta Distribution

$$u(\theta) = \int_0^\theta t^{\alpha-1} \; (1-t)^{\beta-1} \; \partial t$$

$$\frac{\partial}{\partial \tau_s} u(\theta) = \int_0^\theta \left[\log(t) \left(\frac{\partial}{\partial \tau_s} \alpha \right) + \log(1-t) \left(\frac{\partial}{\partial \tau_s} \beta \right) \right]$$

$$* \; t^{\alpha-1} (1-t)^{\beta-1} \partial t$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t} u(\theta) = \int_0^\theta \left(\log(t) \left(\frac{\partial}{\partial \tau_s} \alpha \right) + \log(1-t) \left(\frac{\partial}{\partial \tau_s} \beta \right) \right)$$

$$* \left(\log(t) \left(\frac{\partial}{\partial \tau_t} \alpha \right) + \log(1-t) \left(\frac{\partial}{\partial \tau_t} \beta \right) \right)$$

$$* \; t^{\alpha-1} (1-t)^{\beta-1} \partial t$$

Marginal Parameters: $\alpha = \alpha_i$, $\beta = \sum_{j=1}^{r} \alpha_j - \alpha_i$

$$\frac{\partial}{\partial \tau_s} \alpha = \begin{cases} 1 & \tau_s = \alpha_i \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial}{\partial \tau_s} \beta = \begin{cases} 0 & \tau_s = \alpha_i \\ 1 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t} \alpha = \frac{\partial^2}{\partial \tau_s \partial \tau_t} \beta = 0$$

Conditional Parameters: $\alpha = \alpha_i$, $\beta = \alpha_{j \neq i}$

$$\frac{\partial}{\partial \tau_s} \alpha = \begin{cases} 1 & \tau_s = \alpha_i \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial}{\partial \tau_s} \beta = \begin{cases} 1 & \tau_s = \alpha_j \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2}{\partial \tau_s \partial \tau_t} \alpha = \frac{\partial^2}{\partial \tau_s \partial \tau_t} \beta = 0$$

Section 5. Generalized Beta Distribution

$$u(\theta) = \int_0^{\theta} \frac{\lambda^{\alpha}\left(\frac{t}{1-t}\right)^{\alpha-1}\left(\frac{1}{1-t}\right)^{2}}{\left[1+\lambda\left(\frac{t}{1-t}\right)\right]^{\alpha+\beta}}\,\partial t$$

$$\frac{\partial}{\partial\tau_s}u(\theta) = \int_0^{\theta}\Bigg\{\left[\log(\lambda) + \log\left(\frac{t}{1-t}\right)\right]\left(\frac{\partial}{\partial\tau_s}\alpha\right)$$

$$-\left[\log\left(1+\lambda\left(\frac{t}{1-t}\right)\right)\right]\left(\frac{\partial}{\partial\tau_s}(\alpha+\beta)\right)$$

$$+\left[\left(\frac{\alpha}{\lambda}\right) - \frac{(\alpha+\beta)\left(\frac{t}{1-t}\right)}{\left[1+\lambda\left(\frac{t}{1-t}\right)\right]}\right]\left(\frac{\partial}{\partial\tau_s}\lambda\right)\Bigg\}\lambda^{\alpha}\left(\frac{t}{1-t}\right)^{\alpha-1}\left(\frac{1}{1-t}\right)^{2}\left[1+\lambda\left(\frac{t}{1-t}\right)\right]^{-\alpha-\beta}\partial t$$

$$\frac{\partial^2}{\partial\tau_s\partial\tau_t}u(\theta) = \int_0^{\theta}\Bigg[\Bigg\{\left[\log(\lambda)+\log\left(\frac{t}{1-t}\right)\right]\left(\frac{\partial}{\partial\tau_s}\alpha\right) - \left[\log\left(1+\lambda\left(\frac{t}{1-t}\right)\right)\right]\left(\frac{\partial}{\partial\tau_s}(\alpha+\beta)\right)$$

$$+\left[\left(\frac{\alpha}{\lambda}\right) - (\alpha+\beta)\left(\frac{t}{1-t}\right)\left(1+\lambda\left(\frac{t}{1-t}\right)\right)^{-1}\right]\left(\frac{\partial}{\partial\tau_s}\lambda\right)\Bigg\}$$

$$\star\Bigg\{\left[\log(\lambda)+\log\left(\frac{t}{1-t}\right)\right]\left(\frac{\partial}{\partial\tau_t}\alpha\right)$$

$$-\left[\log\left(1+\lambda\left(\frac{t}{1-t}\right)\right)\right]\left(\frac{\partial}{\partial\tau_t}(\alpha+\beta)\right)$$

$$+\left[\left(\frac{\alpha}{\lambda}\right) - (\alpha+\beta)\left(\frac{t}{1-t}\right)\left[1+\lambda\left(\frac{t}{1-t}\right)\right]^{-1}\right]\left(\frac{\partial}{\partial\tau_t}\lambda\right)\Bigg\}$$

$$+\Bigg\{\left[\left(\frac{1}{\lambda}\right)\left(\frac{\partial}{\partial\tau_t}\alpha\right) - \left(\frac{\alpha}{\lambda^2}\right)\left(\frac{\partial}{\partial\tau_t}\lambda\right) - \frac{\left(\frac{t}{1-t}\right)}{\left[1+\lambda\left(\frac{t}{1-t}\right)\right]}\left(\frac{\partial}{\partial\tau_t}(\alpha+\beta)\right)\right.$$

$$\left.+\,(\alpha+\beta)\left(\frac{t}{1-t}\right)^{2}\left(1+\lambda\left(\frac{t}{1-t}\right)\right)^{-2}\left(\frac{\partial}{\partial\tau_t}\lambda\right)\right]\left(\frac{\partial}{\partial\tau_s}\lambda\right)$$

$$+\left[\left(\frac{\alpha}{\lambda}\right) - (\alpha+\beta)\left(\frac{t}{1-t}\right)\left(1+\lambda\left(\frac{t}{1-t}\right)\right)^{-1}\right]\left(\frac{\partial^2}{\partial\tau_s\partial\tau_t}\lambda\right)\Bigg\}\Bigg]$$

$$\star\left[\lambda^{\alpha}\left(\frac{t}{1-t}\right)^{\alpha-1}\left(\frac{1}{1-t}\right)^{2}\left(1+\lambda\left(\frac{t}{1-t}\right)\right)^{-\alpha-\beta}\right]\partial t$$

Marginal Parameters: $\alpha = \alpha_i$, $\alpha+\beta = \alpha_o + \alpha_i$, $\lambda = \lambda_i$

$$\frac{\partial}{\partial \tau_s} = \begin{cases} 1 & \tau_s = \alpha_i \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial}{\partial \tau_s}(\alpha+\beta) = \begin{cases} 1 & \tau_s = \alpha_o \text{ or } \alpha_i \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial}{\partial \tau_s}\lambda = \begin{cases} 1 & \tau_s = \lambda_i \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2}{\partial \tau_s \, \partial \tau_t} = \frac{\partial^2}{\partial \tau_s \, \partial \tau_t}(\alpha+\beta) = \frac{\partial^2}{\partial \tau_s \, \partial \tau_t}\lambda = 0$$

Conditional Parameters: $\alpha = \theta_i$, $\alpha+\beta = \sum_{j=0}^{r}\theta_j$, $\lambda = \dfrac{\lambda_i}{1+\sum_{j\neq i}\lambda_j\frac{t_j}{1-t_j}}$

$$\frac{\partial}{\partial\theta_s}\alpha = \begin{cases} 1 & \theta_s = \theta_i \\ 0 & \text{otherwise} \end{cases} \qquad \frac{\partial^2}{\partial\theta_s\partial\theta_t}\alpha = 0$$

$$\frac{\partial}{\partial\theta_s}(\alpha+\beta) = \begin{cases} 1 & \theta_s = \theta_j \\ 0 & \text{otherwise} \end{cases} \qquad \frac{\partial^2}{\partial\theta_s\partial\theta_t}(\alpha+\beta) = 0$$

$$\frac{\partial}{\partial\theta_s}\lambda = \begin{cases} \left(1+\sum_{j\neq i}\lambda_j\left(\frac{t_j}{1-t_j}\right)\right)^{-1} & \theta_s = \theta_i \\ -\lambda_i\frac{t_k}{1-t_k}\left(1+\sum_{j\neq i}\lambda_j\left(\frac{t_j}{1-t_j}\right)\right)^{-2} & \theta_s = \theta_{k\neq i} \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial^2}{\partial\theta_s\partial\theta_t}\lambda = \begin{cases} 0 & \theta_s = \theta_t = \theta_i \\ -\left(\frac{t_k}{1-t_k}\right)\left(1+\sum_{j\neq i}\lambda_j\left(\frac{t_j}{1-t_j}\right)\right)^{-2} & \theta_s = \theta_i,\ \theta_t = \theta_{k\neq i} \\ -2\lambda_i\left(\frac{t_k}{1-t_k}\right)^2\left(1+\sum_{j\neq i}\lambda_j\left(\frac{t_j}{1-t_j}\right)\right)^{-3} & \theta_s = \theta_t = \theta_{k\neq i} \\ -2\lambda_i\left(\frac{t_k}{1-t_k}\right)\left(\frac{t_{k'}}{1-t_{k'}}\right)\left(1+\sum_{j\neq i}\lambda_j\frac{t_j}{1-t_j}\right)^{-3} & \theta_s = \theta_{k\neq i},\ \theta_t = \theta_{k'\neq i},\ k\neq k' \\ 0 & \text{otherwise} \end{cases}$$

References

Balch, M. S. "On recent developments in subjective expected utility." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Balch, M. S., and Fishburn, P. C. "Subjective expected utility for conditional primitives." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974

Balch, M. S.; McFadden, D. L.; and Wu S. Y. (eds.). Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Berman, A., and Plemmons, R. J. "Cones and iterative methods for best least squares solutions of linear systems." SIAM Journal of Numerical Analysis, 1974, 11, 145-154.

Bloxom, B. "Individual differences in multidimensional scaling." Research Bulletin 68-45, Princeton, N. J.: Educational Testing Service, 1968.

Bolker, E. D. "A simultaneous axiomatization of utility and subjective probability." Philosophy of Science, 1967, 34, 333-340.

Bolker, E. D. "Remarks on 'Subjective expected utility for conditional primitives'." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Campbell, N. R. Physics: The Elements. Cambridge, England: Cambridge University Press, 1920. Republished as Foundations of Science. New York: Dover Publications, 1957.

Carroll, J. D. "Categorical Conjoint Measurement." Meeting of Mathematical Psychology, August, 1969.

Carroll, J. D. "Spatial, non-spatial, and hybrid models for scaling." Psychometrika, 1976, 41, 439-464.

Carroll, J. D., and Chang, J.-J. "Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart-Young' decomposition." Psychometrika, 1970, 35, 283-319.

Carroll, J. D., and Wish, M. "Models and methods for three-way multidimensional scaling." In D. H. Krantz, N. H. Anderson, R. L. Luce, and P. Suppes, (eds.) Contemporary Developments in Mathematical Psychology, Volume I. San Francisco: W. H. Freeman and Company, 1974.

Céa, J., and Glowinski, R. "Sur des methodes d'optimization par relaxation." Revue Francaise d'Automatique, Informatique, et Researche Opèrationelle, Section R3, 1973, 7, 5-32.

Chen, J. J., and Novick, M. R. "On the use of a cumulative distribution as a utility function in educational or employment selection." Office of Naval Research Technical Report 81-1, Division of Educational Psychology, Measurement, and Statistics, The University of Iowa, Iowa City, Iowa, February, 1981.

Coombs, C. H. "Psychological scaling without a unit of measurement." Psychological Review, 1950, 57, 145-158.

Coombs, C. H. "Mathematical models in psychological scaling." Journal of the American Statistical Association, 1951, 46, 480-489.

Coombs, C. H. A Theory of Data. New York: John Wiley and Sons, 1964.

Coombs, C. H. "Portfolio theory and the measurement of risk." In M. F. Kaplan and S. Schwartz (eds.), Human Judgment and Decision Processes. New York: Academic Press, Inc., 1975.

Dahlquist, G.; Bjorck, A.; and Anderson, N. Numerical Methods. Englewood Cliffs, New Jersey: Prentice Hall, 1974.

Davidson, D.; Suppes, P.; and Siegel, S. Decision Making: An Experimental Approach. Stanford, California: Stanford University Press, 1957.

Dawes, R. M. "Objective optimization under multiple subjective functions." In J. L. Cochrane and M. Zeleny (eds.), Multiple Criteria Decision Making. Columbia, South Carolina: University of South Carolina Press, 1973.

de Finetti, B. Theory of Probability, two volumes. London: John Wiley and Sons, 1974.

de Leeuw, J. "A normalized cone regression approach to alternating least squares algorithms." Unpublished manuscript, Department of Data Theory, University of Leiden, Leiden, The Netherlands, undated.

de Leeuw, J., and Pruzansky, S. "A new computational method to fit the weighted Euclidean distance model." Unpublished manuscript, Bell Laboratory, 1976.

de Leeuw, J.; Young, F. W.; and Takane, Y. "Additive structure in qualitative data: An alternating least squares method with optimal scaling features." Psychometrika, 1976, 41, 471-503.

Debreu, G. "Topological methods in cardinal utility theory." In K. J. Arrow, S. Karlin, and P. Suppes (eds.), Mathematical Methods in the Social Sciences. Stanford, California: Stanford University Press, 1960.

Eckart, C., and Young, G. "Approximation of one matrix by another of lower rank." Psychometrika, 1936, 1, 211-218.

Edwards, W. "Subjective probabilities inferred from decisions." In R. D. Luce, R. R. Bush, and E. Galanter (eds.), Readings in Mathematical Psychology, Volume II. New York: John Wiley and Sons, 1965.

Fishburn, P. C. Decisions and Value Theory. New York: John Wiley and Sons, 1964.

Fishburn, P. C. "Independence in utility theory with whole product sets." Operations Research, 1965, 13,

Fishburn, P. C. "Additivity in utility theory with denumerable product sets." Econometrica, 1966, 34, 500-503.

Fishburn, P. C. "Additive utilities with incomplete product sets: Applications to priorities and assignments." Operations Research, 1967a, 15, 537-542.

Fishburn, P. C. "Conjoint measurement in utility theory with incomplete product sets." Journal of Mathematical Psychology, 1967b, 1, 104-199.

Fishburn, P. C. Utility Theory for Decision Making. New York: John Wiley and Sons, 1970.

Fishburn, P. C. "Approximations of multiattribute utility functions." 1973a.

Fishburn, P. C. "Bernoullian utilities for multiple-factor situations." In J. L. Cochrane and M. Zeleny (eds.), Multiple Criterion Decision Making. Columbia, South Carolina: the University of South Carolina Press, 1973b.

Fishburn, P. C. "A mixture-set axiomatization of conditional subjective expected utility." Econometrica, 1973c, 41, 1-25.

Fishburn, P. C. "On the foundations of decision making under uncertainty." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Fishburn, P. C. "Approximations of two-attribute utility functions." Mathematics of Operations Research, 1977a, 2, 30-44.

Fishburn, P. C. "A survey of multiattribute/multicriterion evaluation theories." 1977b.

Fishburn, P. C. "Two-piece von Neumann-Morgenstern utility functions." Undated.

Fisher, R. A. Statistical Methods for Research Workers (10th edition). Edinburgh: Oliver and Boyd, 1946.

Friedman, M., and Savage, L. J. "The utility analysis of choice involving risk." Journal of Political Economy, 1948, 56, 279-304.

Green, P. E., and Wind, Y. Multiattribute Decisions in Marketing: A Measurement Approach. Hinsdale, Illinois: The Dryden Press, 1973.

Guttman, L. "A general nonmetric technique for finding the smallest coordinate space for a configuration of points." Psychometrika, 1968, 33, 469-506.

Hendrix, W. H.; Ward, J. H., Jr.; Pina, M., Jr.; and Haney, D. L. "Pre-enlistment person-job match system." AFHRL-TR-79-29, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas, 1979.

Herstein, I. N., and Milnor, J. "An axiomatic approach to measurable utility." Econometrica, 1953, 21, 291-297.

Hilgard, E. R. Theories of Learning. New York: Appleton-Century-Croft, 1956.

Horan, C. G. "Multidimensional scaling: Combining observations when individuals have different perceptual structures." Psychometrika, 1969, 34, 139-165.

Hotelling, H. "Analysis of a complex of statistical variables into principle components." Journal of Educational Psychology, 1938, 24, 417-441.

Humphreys, P., and Wisudha, A. "Multi Attribute Utility Decomposition." Technical Report 79-2, Brunel University, Uxbridge, Middlesex, England, April, 1979.

Isaacs, G. L., and Novick, M. R. Manual for the Computer-Assisted Data Analysis Monitor-1978. Iowa City, Iowa: The University of Iowa, 1978.

Jaffray, J.-Y. "On the extension of additive utilities to infinite sets." Journal of Mathematical Psychology, 1974, 11, 430-458.

Jeffrey, R. C. "New foundations for Bayesian decision theory." In Y. Bar-Hillel (ed.), Proceedings of the 1964 International Congress on Logic, Methodology and Philosophy of Science. Amsterdam: North-Holland Publishing Company, 1965a.

Jeffrey, R. C. The Logic of Decision. New York: McGraw-Hill Book Company, 1965b.

Jeffrey, R. C. "Frameworks for preference." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Johnson, N. L., and Kotz, S. Distributions in Statistics: Continuous Multivariate Distributions. New York: John Wiley and Sons, 1972.

Kahnemann, D., and Tversky, A. "Prospect theory: An analysis of decision under risk." Econometrica, 1979, 47, 263-291.

Keeney, R. L. "Utility independence and preference for multiattributed consequences." Operations Research, 1971, 19, 875-893.

Keeney, R. L. "Utility functions for multiattributed consequences." *Management Science*, 1972, *18*, 276-287.

Keeney, R. L., and Raiffa, H. *Decisions With Multiple Objectives: Preferences and Value Trade-offs*. New York: John Wiley and Sons, 1976.

Keeney, R. L., and Sicherman, A. "An interactive computer program for assessing and analyzing preferences concerning multiple objectives." *Research Memorandum RM-75-2*, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1975.

Kennedy, W. J., Jr., and Gentle, J. E. *Statistical Computing*. New York: Marcel Dekkar, 1980.

Krantz, D. H. "The Luce-Tukey axiomatization and some extension." *Journal of Mathematical Psychology*, 1964, *1*, 248-277.

Krantz, D. H., and Luce, R. D. "The interpretation of conditional expected-utility theories." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), *Essays on Economic Behavior Under Uncertainty*. Amsterdam: North-Holland Publishing Company, 1974.

Krantz, D. H.; Luce, R. D.; Suppes, P.; and Tversky, A. *Foundations of Measurement*. Volumes I and II, New York: Academic Press, 1972.

Kruskal, J. B. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis." *Psychometrika*, 1964, *29*, 1-227.

Kruskal, J. B. "Analysis of factorial experiments by estimating monotone transformations of the data." *Journal of the Royal Statistical Society*, Series B, 1965, *27*, 251-263.

Kruskal, J. B., and Shepard, R. N. "A nonmetric variety of linear factor analysis." *Psychometrika*, 1974, *39*, 123-157.

Lee, S. Y., and Jennrich, R. I. "A study of algorithms for covariance structure analysis with specific comparisons using factor analysis." *Psychometrika*, 1979, *44*, 99-113.

Libby, D. L. "Factor Analysis as Multidimensional Scaling." Unpublished, 1979.

Lingoes, J. C. "A general survey of the Guttman-Lingoes nonmetric program series." In R. N. Shepard, A. K. Romney, and S. B. Nerlove (eds.), *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, Volume I. New York: Seminar Press, 1972.

Lingoes, J. C., and Guttman, L. "Nonmetric factor analysis: A rank reducing alternative to linear factor analysis." *Multivariate Behavioral Research*, 1967, *12*, 485-505.

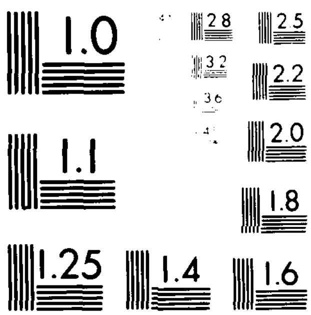
Luce, R. D. *Individual Choice Behavior*. New York: John Wiley and Sons, 1959.

Luce, R. D. "Two extensions of conjoint measurement." *Journal of Mathematical Psychology*, 1966a, *3*, 348-370.

AD-A098 230 IOWA UNIV IOWA CITY DIV OF EDUCATIONAL PSYCHOLOGY F/G 12/1
MULTIATTRIBUTE FIXED-STATE UTILITY ASSESSMENT.(U)
MAR 81 D L LIBBY, M R NOVICK N00014-77-C-0428
UNCLASSIFIED TR-81-4 NL

3 of 3
AD A
098230

END
DATE
FILMED
5-81
DTIC



MICROCOPY RESOLUTION TEST CHART

Luce, R. D. "A 'fundamental' axiomatization of multiplicative power relationships among three variables." Philosophy of Science, 1966b, 32.

Luce, R. D., and Krantz, D. H. "Conditional expected utilities." Econometrica, 1971, 39, 253-271.

Luce, R. D., and Raiffa, H. Games and Decisions. New York: John Wiley and Sons, 1957.

Luce, R. D., and Tukey, J. W. "Simultaneous conjoint measurement: A new type of fundamental measurement." Journal of Mathematical Psychology, 1964, 1, 1-27.

Mayekawa, S.-I. "Note on the possibility of nonmetric scaling of utility." Unpublished, 1980.

Mayekawa, S.-I. "Fixed-state utility scaling: Comparison of several criteria." Office of Naval Research Technical Report 81-3, Division of Educational Psychology, Measurement, and Statisitcs, The University of Iowa, Iowa City, Iowa, February, 1981.

Mosteller, F., and Nogee, P. "An experimental measurement of utility." Journal of Political Economy, 1951, 59, 371-404.

Narens, L. "Minimal conditions for additive conjoint measurement and qualitative probability." Journal of Mathematical Psychology, 1974, 11, 404-430.

Novick, M. R. "Statistics as Psychometrics." Psychometrika, 1980, 45, 411-424.

Novick, M. R.; Chuang, D.; and DeKeyrel, D. F. "Local and regional coherence utility assessment procedures." Trabajos de Estadistica, (in press).

Novick, M. R.; Hamer, R. M.; Libby, D. L.; Chen, J. J.; and Woodworth, G. G. The Computer-Assisted Data Analysis (CADA) Monitor. Iowa City, Iowa: The University of Iowa, 1980.

Novick, M. R., and Jackson, P. H. Statistical Methods for Educational and Psychological Research. New York: McGraw Hill Book Company, 1974.

Novick, M. R., and Lindley, D. V. "The use of more realistic utility functions in educational applications." Journal of Educational Measurement, 1978, 15, 181-192.

Novick, M. R., and Lindley, D. V. "Fixed state assessment of utility functions." Journal of the American Statistical Association, 1979, 74, 306-311.

Novick, M. R.; Turner, N.; and Novick, L. R. "Experimental studies of CADA-based utility assessment procedures." Office of Naval Research Technical Report 81-4, Division of Educational Psychology, Measurement, and Statistics, The University of Iowa, Iowa City, Iowa, March, 1981.

Pollak, R. A. "Additive von Neumann-Morgenstern utility functions." Econometrika, 1967, 35, 485-595.

Pratt, J. W. "Some comments on some axioms for decision making under uncertainty." In M. S. Balch, D. L. McFadden, and S. Y. Wu (eds.), Essays on Economic Behavior Under Uncertainty. Amsterdam: North-Holland Publishing Company, 1974.

Pratt, J. W.; Raiffa, H.; and Schlaifer, R. O. Introduction to Statistical Decision Theory. New York: McGraw-Hill Book Company, 1965.

Raiffa, H., and Schlaifer, R. O. Applied Statistical Decision Theory. Cambridge, Mass.: Harvard Business School Press, 1961.

Ramsey, F. P. The Foundations of Mathematics and Other Logical Essays. Paterson, New Jersey: Littlefield, Adams and Co., 1960.

Roskies, R. A. "A measurement axiomatization for an essentially multiplicative representation of two factors." Journal of Mathematical Psychology, 1965, 2, 266-276.

Savage, L. J. The Foundations of Statistics. New York: John Wiley and Sons, 1954.

Schlaifer, R. O. Probability and Statistics for Business Decisions: An Introduction to Managerial Economics Under Uncertainty. New York: McGraw-Hill Book Company, 1959.

Schlaifer, R. O. Analysis of Decisions Under Uncertainty. New York: McGraw-Hill Book Company, 1969.

Schoneman, P. H. "An algebraic solution for a class of subjective metrics models." Psychometrika, 1972, 37, 441-451.

Shepard, R. N. "Metric structures in ordinal data." Journal of Mathematical Psychology, 1966, 3, 287-315.

Shepard, R. N., and Kruskal, J. B. "Nonmetric methods for scaling and for factor analysis." American Psychologist, 1964, 19, 557-558.

Shepard, R. N.; Romney, A. K.; and Nerlove, S. E. "Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Volume I. New York: Seminar Press, 1972.

Siegel, S. Nonparametric Statistics. New York: McGraw-Hill Book Company, 1956a.

Siegel, S. "A method for obtaining an ordered metric scale." Psychometrika, 1956b, 21, 207-216.

Slovic, P. "From Shakespeare to Simon: Speculation -- and some evidence -- about man's ability to process information." Oregon Research Institute Research Bulletin, 1972, 12, 2.

Spaeth, H. J., and Guthery, S. B. "The use and utility of the monotone criterion in multidimensional scaling." Multivariate Behavioral Research, 1969, 4, 501-515.

Spetzler, C. S., and Stael von Holstein, C-A. "Probability encoding in decision analysis." Management Science, 1975, 22, 340-358.

Stevens, S. S. "On the theory of scales of measurement." Science, 1946, 103, 677-680.

Stevens, S. S. "Mathematics, measurement, and psychophysics." In S. S. Stevens (ed.), Handbook of Experimental Psychology. New York: John Wiley and Sons, 1951.

Suppes, P., and Winet, M. "An axiomatization of utility based on the notion of utility differences." Management Science, 1955, 1, 259-270.

Swalm, R. O. "Utility theory -- insights into risk taking." Harvard Business Review, 1966, 44, 123-136.

Takane, Y.; Young, F. W.; and de Leeuw, J. "Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features." Psychometrika, 1977, 42, 7-67.

Torgerson, W. S. "Multidimensional scaling: I Theory and method." Psychometrika, 1952, 17, 401-419.

Torgerson, W. S. Theory and Methods of Scaling. New York: John Wiley and Sons, 1960.

Tucker, L. R., and Messick, S. "An individual differences model for multidimensional scaling." Psychometrika, 1963, 28, 333-367.

Tversky, A. "A general theory of polynomial conjoint measurement." Journal of Mathematical Psychology, 1967a, 4, 1-20.

Tversky, A. "Additivity, utility, and subjective probability." Journal of Mathematical Psychology, 1967b, 4, 175-201.

Tversky, A. "On the elicitation of preferences: Descriptive and prescriptive considerations." In D. E. Bell, R. L. Keeney, and H. Raiffa (eds.), Objectives in Decisions. Chichester: John Wiley and Sons, 1977.

Tversky, A., and Kahnemann, D. "Judgment under uncertainty: Heuristics and biases." Science, 1974, 185, 1124-1131.

Tversky, A., and Zivian, A. A. "A computer program for additivity analysis." Journal of Mathematical Psychology, 1966, 11, 78-79.

Ward, J. H., Jr. "Creating mathematical models of judgment processes: from policy-capturing to policy-specifying." AFHRL-TR-77-47, Air Force Human Resources Laboratory, Brooks Air Force Base, Texas, 1977.

Ward, J. H., Jr.; Pina, M., Jr.; Fast, J. C.; and Roberts, D. K. "Policying specifying with applications to personnel classification amd assignment." Air Force Human Resources Laboratory, Brooks Air Force Base, Texas, 1979.

Young, F. W. "A model of polynomial conjoint measurement analysis algorithms." In R. N. Shepard, A. K. Romney, and S. B. Nerlove (eds.), "Multidimensional Scaling:

Theory and Applications in the Behavioral Sciences, Volume I. New York: Seminar Press, 1972.

Young, F. W. "Methods for describing ordinal data with cardinal models." Journal of Mathematical Psychology, 1975, 12, 416-436.

Young, F. W.; de Leeuw, J.; and Takane, Y. "Quantifying qualitative data." Report Number 149. The L. L. Thurstone Psychometric Laboratory, the University of North Carolina, 1976a.

Young, F. W.; de Leeuw, J.; and Takane, Y. "Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features." Psychometrika, 1976b, 41, 505-529.

Young, F. W.; Takane, Y.; and de Leeuw, J. "The principle components of mixed measurement level multivariate data: An alternating least squares method with optimal scaling features." Psychometrika, 1978, 43, 279-281.

Young, G., and Householder, H. S. "Discussion of a set of points in terms of their mutual distances." Psychometrika, 1938, 3, 19-22.

von Neumann, J., and Morgenstern, O. Theory of Games and Economic Behavior. 2nd Edition, Princeton, N. J.: Princeton University Press, 1947.

Navy

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Chief of Naval Education and Training
Liason Office
Air Force Human Resource Laboratory
Flying Training Division
WILLIAMS AFB, AZ 85224

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Henry M. Halff
Department of Psychology,C-009
University of California at San Diego
La Jolla, CA 92093

1 Dr. Patrick R. Harrison
Psychology Course Director
LEADERSHIP & LAW DEPT. (7b)
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

Navy

1 CDR Robert S. Kennedy
Head, Human Performance Sciences
Naval Aerospace Medical Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. Kneale Marshall
Scientific Advisor to DCNO(MPT)
OP01T
Washington DC 20370

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Navy

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
(OP-115)
Washington, DC 20350

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

1 Dr. Bernard Rimland (03B)
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

Navy

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Ronald Weitzman
Code 54 WZ
Department of Administrative Sciences
U. S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Myron Fischl
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Dexter Fletcher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria,VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Army

1 Commandant
US Army Institute of Administration
Attn: Dr. Sherrill
FT Benjamin Harrison, IN 46256

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Research and Measurment Division
Research Branch, AFMPC/MPCYPR
Randolph AFB, TX 78148

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Dr. Marty Rockway
Technical Director
AFHRL(OT)
Williams AFB, AZ 58224

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

Coast Guard

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Other DoD

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Dr. William Graham
Testing Directorate
MEPCOM/MEPCT-P
Ft. Sheridan, IL 60037

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 MAJOR Wayne Sellman, USAF
Office of the Assistant Secretary
of Defense (MRA&L)
3B930 The Pentagon
Washington, DC 20301

1 DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civil Govt

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Werner Birke
DezWPs im Streitkraefteamt
Postfach 20 50 03
D-5300 Bonn 2
WEST GERMANY

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52240

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

Non Govt

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. William E. Coffman
Director, Iowa Testing Programs
334 Lindquist Center
University of Iowa
Iowa City, IA 52242

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

1 Dr. Leonard Feldt
Lindquist Center for Measurment
University of Iowa
Iowa City, IA 52242

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Non Govt

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
Suite 900
4330 East West Highway
Washington, DC 20014

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Chester Harris
School of Education
University of California
Santa Barbara, CA 93106

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Non Govt

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Professor John A. Keats
University of Newcastle
AUSTRALIA 2308

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. Gary Marco
Educational Testing Service
Princeton, NJ 08450

1 Dr. Scott Maxwell
Department of Psychology
University of Houston
Houston, TX 77004

Non Govt

1 Dr. Samuel T. Mayo
Loyola University of Chicago
820 North Michigan Avenue
Chicago, IL 60611

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
4 Hill Hall
Columbia, MO 65211

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

Non Govt

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Kazuo Shigemasu
University of Tohoku
Department of Educational Psychology
Kawauchi, Sendai 980
JAPAN

1 Dr. Edwin Shirkey
Department of Psychology
University of Central Florida
Orlando, FL 32816

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

Non Govt

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN
THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Brad Sympson
Psychometric Research Group
Educational Testing Service
Princeton, NJ 08541

1 Dr. Kikumi Tatsuoka
Computer Based Education Research
Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Non Govt

1 Dr. Howard Wainer
Bureau of Social SCience Research
1990 M Street, N. W.
Washington, DC 20036

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

1 Wolfgang Wildgrube
Streitkraefteamt
Box 20 50 03
D-5300 Bonn 2

DATE

FILMED

-8